The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

The role of Set1 methylation of histone H3 lysine 4 on chromatin structure in *Saccharomyces cerevisiae*

Rochan Mars

University of Cape Town

SUBMITTED IN FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN THE DEPARTMENT OF MOLECULAR AND CELL BIOLOGY
FACULTY OF SCIENCE
UNIVERSITY OF CAPE TOWN

2005

CONTENTS

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

Abstract

It has been established that the *S .cerevisiae* Set1p is a histone H3 methyltransferase and in its absence causes a decrease in transcriptional activity of approximately 80% of ORFs (Boa, *et al.*, 2003). To understand the role of histone H3 lysine 4 methylation and how it influenced chromatin structure, nuclei from wildtype and *set1*⁻ strains were isolated and treated with different concentrations of micrococcal nuclease (Mnase). It was found that the wildtype nuclei were more accessible to Mnase in comparison to the *set1*⁻ strain, which had the appearance of heterochromatin and a smaller apparent repeat length. The role of Set1p H3K4 methylation on chromatin structure was analyzed at high resolution at the *SIR4* and *TRX2* promoters. The *Sir4* gene was previously shown to be upregulated and the *TRX2* gene down-regulated in the *set1*⁻ strain. The internucleosomal DNA at the *SIR4* promoter in the *set1*⁻ strain was more accessible and more extensively cleaved than that of the wildtype strain. This suggested that the nucleosomes were differently positioned on the *SIR4* promoter in the *set1*⁻ and wildtype strains. The nucleosomes on the *TRX2* promoter had footprints that were shifted in the *set1*⁻ strain in comparison to the wildtype, and showed greater accessibility than the wildtype. These results supported the idea that histone H3 lysine 4 methylation indirectly influenced chromatin structure.

University of Cape Town

ABBREVIATIONS

°C	Degrees celsius
Å	angstrom
A_{260}	Absorbance at 260nm
bp	base pair
C-terminus	Carboxy terminus
Da	Dalton
dCTP	deoxycytosine triphosphate
dNTP	deoxynucleotide triphosphate
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylene-diamene-tetra-acetic acid
E(var)	Enhancer of varigation
g	gram or gravity constant
GC-rich	DNA rich in guanine and cytosine
h	hour
H_2O	water
His	L-Histidine
kb	kilobases
kDa	kilodalton
mg	milligram
min	minute
µg	microgram
µM	micromolar

University of Cape Town

µmol	micromole
MDa	Mega Dalton
Mnase	micrococcal nuclease
mm	millimetre
mM	millimolar
mmol	millimole
mol	mole
nm	nanometre
N-terminus	Amino-terminus
PcG	Polycomb group
OD_{600}	Optical Density at 600 nm
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
RNA	Ribonucleic acid
mRNA	Messenger ribonucleic acid
Hho1	Hho1 protein
rRNA	Ribosomal ribonucleic acid
rDNA	Ribosomal deoxyribonucleic Acid
rpm	Revolutions per minute
s	second (unit of time)
SDS	sodium dodecyl sulphate
Su(var)	Suppressor of variegation
TrxG	Trithorax group
Tris	Tris(hydroxymethyl)aminomethane
U	unit

University of Cape Town

UV	Ultraviolet
V	Volt
v/v	volume per volume
w/v	weight per volume

University of Cape Town

LIST OF FIGURES

University of Cape Town

University of Cape Town

University of Cape Town

LIST OF TABLES

University of Cape Town

ACKNOWLEDGEMENTS

I would like to thank the Lord God for the strength and guidance He has given me throughout my studies. I am extremely grateful to my supervisor Hugh G. Patterton for giving me the opportunity to be able to part take in chromatin research. I would like to acknowledge the Wellcome Trust and the National Research Foundation for providing the funding which made it possible for our lab to participate in cutting edge molecular biological research. Thank you too, Maureen and Desmond for the autoclaving as well as ensuring that we never ran out of clean glassware. I wish to thank Rosetta and Peter for a job well done. My gratitude goes out to my parents and sisters for their encouragement and support for which no amount of words can describe. I wish to thank my closest friends Waseemah, Zaheer and Randall for standing by me through the good as well as the bad times and whose support meant a lot to me. Last but not least to my lab mates Sachin, Tim, Chris, Lieschen, Georgia and Tony for their advice and support and to Michelline Sanderson for her technical advice.

University of Cape Town

Chapter 1: Introduction

Chromatin structure plays an important role in the packaging of DNA within the eukaryotic nucleus. The structure consists of DNA wrapped around the histone octamer which forms the nucleosome (Zhang *et al.*, 2004). Once thought of as being static complexes involved in higher order chromatin structure, the nucleosomes are now appreciated to be templates for many cellular processes, such as transcription, replication, repair, recombination and chromosome segregation (Rice and Allis, 2001).

1.1 The Histones

The nucleosome core particle consist of two copies of histones H2A, H2B, H3 and H4 that form an octamer wrapped by 146bp of DNA (Wolffe, 1998 and Luger *et al.*, 1997) The average molecular weight of one of these core histones is approximately 14,000Da and consist of between 102-129 amino acids . The linker histone H1 protects an addition length of approximately 20bp in the nucleosome core (Mathews and van Holde, 1996) . H1 has a molecular weight of approximately 23,000Da and consists of 244 amino acids. All core histones share an evolutionary conserved structural motif, which is three alpha helices separated by two loops (Fig 1) (Luger *et al.*, 1997).

University of Cape Town

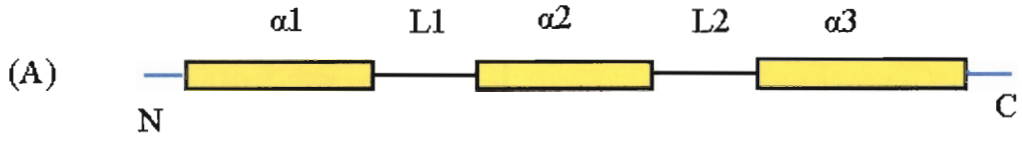


Figure 1: The structural motifs of the core histone proteins

(A) The core histone structure consisting of the alpha 1(α1)-loop1 (L1) - alpha 2(α2)-loop2 (L2) - alpha 3(α3) motif (Luger *et al.*, 1997).

The structure of H1 comprises an unfolded N-terminal region, a central globular region and an extended C-terminal domain (Fig 2). The globular domain is 80 amino acids in length, with approximately 70% of the amino acids in the unfolded N- and C-terminal domains. The yeast *Saccharomyces cerevisiae* linker histone Hho1 has a less conserved structural motif than higher eukaryotes and consists of an unfolded N-terminal region and two globular domains separated by a lysine-rich linker, which appears to be the functional equivalent of the metazoan C-terminal tail (Ono *et al.*, 2003).

University of Cape Town

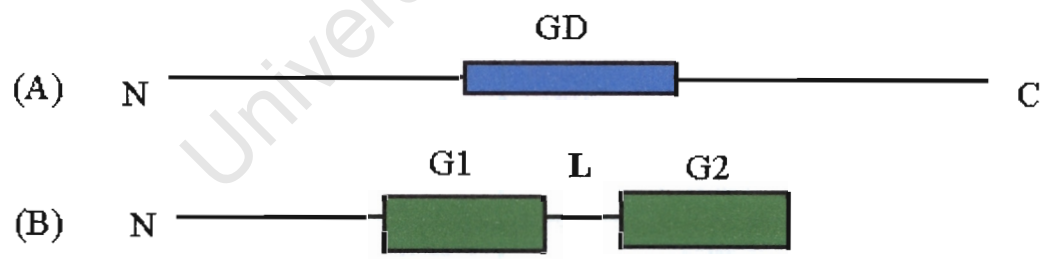


Figure 2: The tertiary structures of the histone H1 variants

(A) H1 structure consisting of an unstructured N-terminus, a globular domain (GD) and a C-terminus **(B)** *S. cerevisiae* Hho1 consisting of the unstructured N-terminus and two globular domains (G1and G2) separated by a linker (L) (Ono *et al.*, 2003).

The histones form a well conserved class of proteins, with histones H3 and H4 being highly conserved and H2A, H2B and H1 showing less sequence conservation (Wolffe, 1998). The human histone H4 has only two amino acid substitutions compared to its orthologue in peas, and there are only eight amino acid substitutions between humans and yeast H4 (Mathews and van Holde, 1996). The N-terminal tails of histones contain a high concentration of basic amino acids such as lysine and arginine. Histone H1 consists of a high percentage of lysine residues in comparison to the core histones, which have an almost equal lysine/arginine ratio (van Holde, 1989). The high concentration of lysine residues within H1 tails is important for higher order chromatin structure, and aids in packaging the nucleosomes into a 30nm fibre, consisting of six nucleosomes per turn, forming a solenoidal structure (Zhang *et al.*, 2001).

University of Cape Town

1.1.1 Histone Structure in the Nucleosome

The crystal structure of the nucleosome core particle at 7Å resolution provided the first insight into the organization of the histone octamer (Richmond *et al.*, 1984). The determination of the crystal structure of the nucleosome core particle was made possible by the isolation of nucleosomes by the micrococcal nuclease digestion of bulk chromatin, which liberated the nucleosome containing a fixed length of DNA. (Simpson, 1978; Dutnall and Ramakrishnan, 1997). The fixed length of DNA ensured that the nucleosomes were homogenous, facilitating crystallization.

Many technical improvements were made to obtain a crystal structure which diffracted at a higher resolution. One of these was the reconstitution of the histone octamer on to the 5S rRNA gene of the sea urchin *Lytechinus variegates,* which is a strong nucleosome

positioning sequence (Richmond *et al.*, 1988).The crystal structure of the nucleosome core particle, solved at 2.8 Å resolution, (Luger *et al.,* 1997) used a palindromic human alpha satellite DNA sequence, which diffracted at high resolution (Dutnall and Ramakrishnan, 1997) (Fig. 3). The use of a palindromic DNA sequence ensured that the nucleosome had two fold symmetry.

At SHL0 (super helix position zero), the nucleosome has a pseudo-two fold axis, dividing the DNA into two 73bp halves, with one base pair exactly aligned with the dyad (Luger *et al.*, 1997) (Fig. 3). The DNA surrounding the histone octamer total 146bp, which is equivalent to 1.65 turns of the left handed super helix. The nucleosome core DNA has a radius of approximately 43Å, a super helical pitch of 27Å and a helical periodicity of 10.2 bp. In contrast, a 147bp DNA nucleosome core particle, solved to 1.9 Å resolution, had a radius of 42Å, a pitch of 26Å and a 10.3 bp helical periodicity. The reasons for these differences were suggested to be differences in the DNA stretching in the core termini (Davey and Richmond, 2003). It was, therefore, evident that the DNA sequence was important in the formation of a compact nucleosomal structure, irrespective of the sequence length.

University of Cape Town

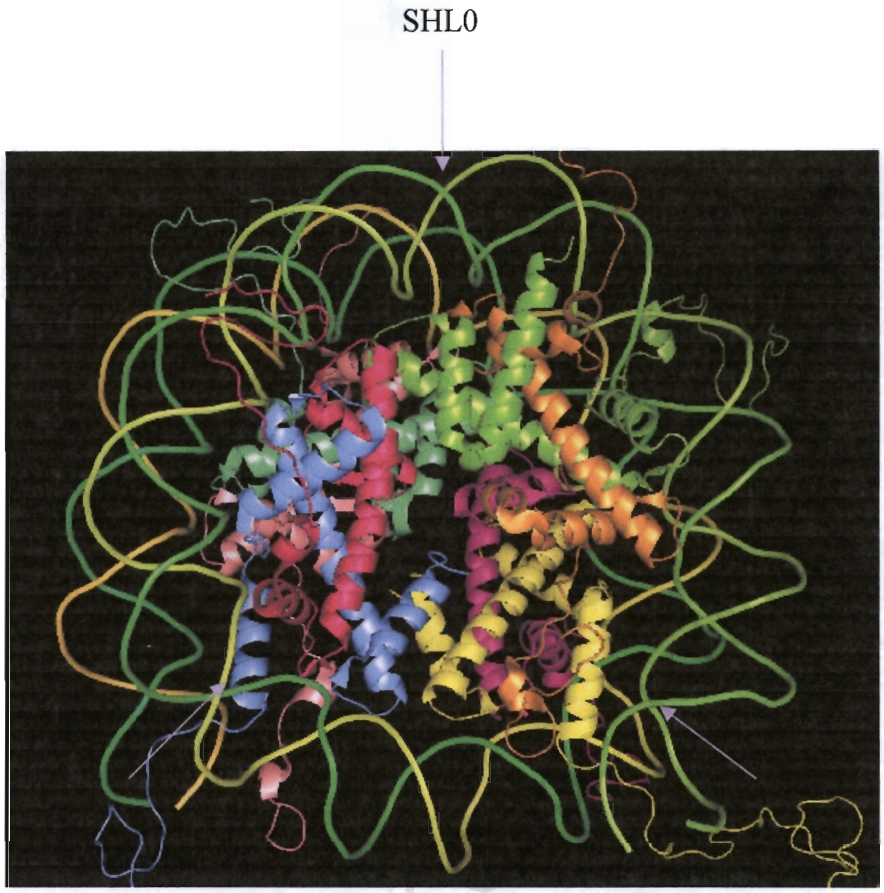


Figure 3: The crystal structure of the nucleosome core particle

The histone octamer consisting of histones H2A (orange and green α helices), H2B (brown and green α helices), H3 (blue and yellow α helices) and H4 (pink and purple α helices) wrapped by 146 bp of DNA. The super helix location (SHL) represents the number of DNA helix turns, starting from SHL0 to SHL7. At SHL0 the nucleosome is divided into 73 and 72bp halves with one base pair situated on the dyad (Luger *et al.*, 1997; Chantalat *et al.*, 2003).

University of

1.2 Nucleosome octamer structure

The histone octamer consists of two histone dimers, (H2A-H2B) $_2$ and one tetramer, (H4-H3) (H3'-H4') (Fig 3) (Chantalat *et al.*, 2003). The two heterodimers (H4-H3) (H3'-H4') interacted through a "four helix bundle" between H3 and H3'. The H2A-H2B dimer interacts with the (H3-H4) tetramer through a homologous four helix bundle between H2B and H4 (Luger *et al.*, 1997; Chantalat *et al.*, 2003). This ensured that the octamer symmetry is two fold, with approximately 121bp of DNA in B-form wrapped onto it (Chantalat *et al.*, 2003; Davey and Richmond, 2003).

1.2.1 DNA –histone interactions

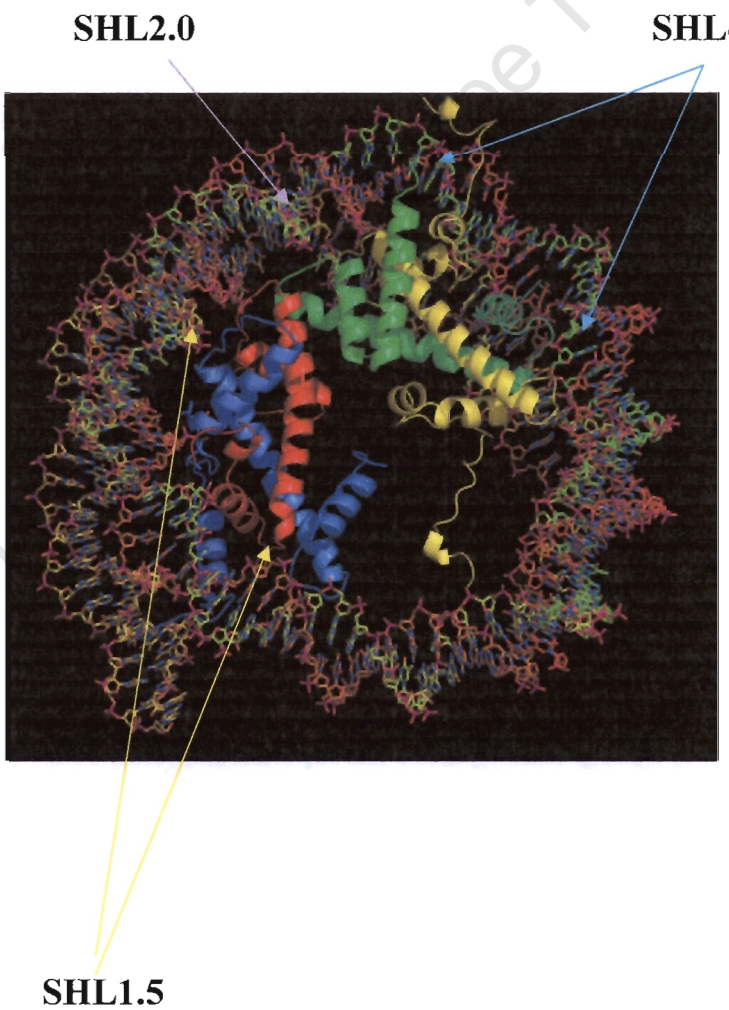
The first high resolution nucleosome core particle structure showed that there were five different features of DNA-histone interactions (Luger *et al.*, 1997). The DNA-histones had two different types of histone binding sites. The first is a α1α1 site, which uses the α1 of H3, H4 and H2B, and α2 of all four core histones to fix the phosphates by a helix dipole charge. The second type comprises the L1 and L2 loops. Hydrogen bonds are formed from the main chain amides in last turn of α1 or α2 of individual histones and through the amide nitrogen atoms from amino acid residues and through DNA phosphate atoms. There are no polar contacts between the DNA and the histones and hydrogen bonding occurs between basic or hydroxyl amino acids. The arginine side chain from the loop L1 chain enters minor groove 10 out 14 times. The H4 loop (L1) arginine is inserted into the DNA minor groove at SHL0.5, for H2A at SHL2.5, for H3 and at SHL5.5 for H2B (Fig 4A) (Luger *et al.*, 1997).

University of Cape Town

The N–terminal tail of H3 passes between the DNA gyres at SHL 6.5-7 and SHL-0.5-0.1. The N-terminal of H2B passes through the DNA gyres at SHL 4.5 and SHL-2.5 to- 3.0 (Fig 4B). These histone N-terminal tails are thought to play a part in higher order chromatin structure such as the 30nm fibre (Dutnall and Ramakrishnan, 1997). The histone N-terminal tails have been shown to be covalently modified, which are thought to alter the chromatin structure (Van Holde, 1989). It has been proposed that the N-terminal covalent modifications constitute a "histone code."

University of Cape Town

A



B

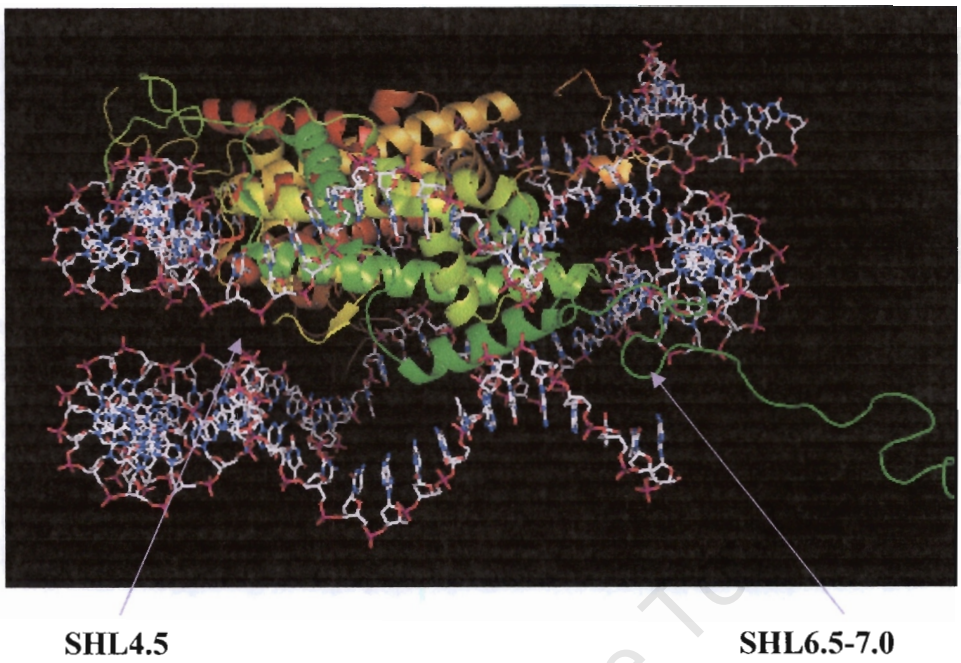


Figure 4: The core histone binding sites

(**A**) The interaction of histones H3 (blue ribbon) and H4 (red ribbon) and their DNA binding sites at SHL1.5 yellow arrows. The histone H2B (yellow ribbon) binds the DNA at SHL4.5 (blue arrows). The histone H2A (green ribbon) interacts with H4 at SHL2.0 (purple arrow). (**B**) The histone tails and their interaction with the DNA super helix (Fig3B) (Luger *et al.*, 1997; *al.*, Chantalat *et al.*, 2003)

University of Cape Town

1.3 Euchromatin and Heterochromatin

Experiments by cytogeneticists revealed that chromatin consisted of two distinct structural states: namely euchromatin and heterochromatin. Euchromatin is DNA-protein fibers that occupy most of the nuclear region, and are not as condensed as mitotic chromosomes. On the other hand, heterochromatin appeared as highly condensed clusters near the nucleolus and nuclear membrane (Lewin 1996; Dillon and Festenstein, 2002).

Heterochromatin can either be constitutive or facultative. Constitutive heterochromatin is found in the pericentromeric region that consisted of repetitive DNA sequences, such as short tandem repeats (STR's) and variable nucleotide tandem repeats (VNTR's) (Dillon and Festenstein, 2002). Facultative heterochromatin is an inactive maternal or paternal X chromosome in a particular cell lineage (Lewin, 1996). The epigenetic transcriptional states of eukaryotic genes are dependent on the position of the gene(s) within an accessible (euchromatic) or inaccessible (heterochromatic) environment, a phenomenom known as position-effect variegation (PEV) (Berger, 2001).

1.4 Suppressors and enhancers of variegation

Cellular differentiation occurs early in development. Cells then maintain their particular state throughout the life-cycle of the organism, without further differentiation (Beuchle *et al.*, 2001).The fruit fly *Drosophila melanogaster* has been used as a model to study the developmental processes of eukaryotes. From early genetic screens, it was determined that segmental development of the head, thorax, and abdomen was accomplished by the transcriptional activation of particular combinations of homeo-box-containing (Hox) selector genes (Beuchle *et al.*, 20001 and Birve *et al.*, 2001). The Hox genes are transcription factors and their particular combinations of activation are important in controlling the development of each segment. Hox genes have the ability to be active both within and outside their correct expression domains but are repressed in cells where they have to remain inactive. The propagation of the heritable silenced state of the Hox genes within *Drosophila* occurred in two steps. Firstly, during embryogenesis Hox gene silencing depended upon the expression of the Gap gene products within that particular

University of Cape Town

cell line. These proteins, including Hunchback (Hb), Krüppel (Kr) and Knirps (kni), bind to *cis*-acting regulatory sequences of the Hox genes, which then repressed transcription (Beuchle *et al.*, 2001). During the remainder of the organism's development (in the absence of the Gap gene products), the propagation of the heterochromatic (inactive) Hox state required the products of the Polycomb (PcG) genes (Beuchle *et al.*, 2001).

The regulation of the Hox genes during *Drosophila* embryogenesis is abolished when there are mutations in PcG genes. These lacks- of- function PcG mutants were unable to prevent the expression of Hox genes, and this led to Hox gene expression outside of its spatial domain. A mutation in 12 PcG genes was shown to be involved in the untimely gene expression of Hox genes during *Drosophila* development (Birve *et al.*, 2001 and Beuchle *et al.*, 2001). Hox genes have specific *cis*-regulatory sequences known as Polycomb response elements (PREs), which regulate their expression by the binding of PcG proteins to nucleosomes positioned on these DNA regulatory elements. Elimination of these *cis*-regulatory elements caused expression of these Hox genes.

The PcG proteins form part of two multimeric complexes, namely the 1-2MDa PRC 1 complex, and the 600 kDa ESC-E (z) complex (Simon and Tamkun, 2002; Müller *et al.*, 2002). There is another group of proteins known as the trithorax group (trxG), which counteracts the effects of the PcG proteins. These proteins are not transcription activators, but are responsible for maintaining the active gene expression of the HOM-C genes within their spatial boundaries (La Jeunesse and Shearn, 1996). The trx genes were identified in the same manner as the PcG genes, namely by screening of mutants that showed similar phenotypes to that of Hox gene mutants. The trxG proteins have been shown to function as four distinct complexes in *Drosophila* (Müller *et al.*, 2002). These

University of Cape Town

complexes are: a 2MDa BRM complex, 2MDa ASH1 complex, a 500kDa ASH2 complex and a 1MDa TRX complex (TAC1) (Simon and Tamkun, 2002; Müller *et al.*, 2002).

The PcG and trxG proteins can be divided into two groups: one which suppresses variegation [the Su (var) group], containing the PcG proteins, and the other that enhances variegation [E (var)], and that contains the trxG proteins. These [Su (var)] and [E (var)] group proteins have been shown to encode protein modules which were associated with histone post-translational modifications, chromatin remodeling and silencing complexes (Jenuwein and Allis, 2001).

1.5 Post-translational modifications

The histones remain an integral part in maintaining higher order chromatin structure and are key components in regulating the machinery responsible for transcription (Strahl and Allis, 2001). As mentioned, the histone tails are subjected to various post-translational modifications (Table 1), which occur on all core histones. These covalent modifications are hypothesized to alter chromatin structure, thus allowing the unmasking of DNA regulating elements to various transcription factors (Berger, 2001). Extensive experimental evidence suggested that different histone modifications were linked to various transcriptional states, which are either "euchromatic" (on) or "heterochromatic" (off) (Jenuwein and Allis, 2001). As mentioned, the PcG proteins which formed part of the sub-class of the Su [var] group were involved in maintaining the heterochromatic state of genes. The Su[var] genes coded for enzymes and chromatin associated proteins such as histone deacetylases (HDACs), protein phosphatases (PPTases), S-adenosylmethionine (SAM) synthetase, the heterochromatin associated protein HP1

University of Cape Town

[Su(var) 2-5]. These chromatin associated proteins were found to posses various domains that may modulate chromatin structure, including the bromodomain which is characteristic of histone acetyltransferases (HATs), SNF2, TAF_{II} 250 and the mammalian trithorax (HRX/MLL) and was shown to bind to acetylated lysine residues in histone H3 (Jenuwein and Allis, 2001). The chromo domain is characteristic of proteins involved in heterochromatin formation, such as the Polycomb group and HP1 proteins. The histone methyltransferases (HMTs), with the exception of the *disruptor of telomeric silencing protein* (Dot1p), has the characteristic SET domain, which is found in *Su* (var) 3-9, *E* (z) a member of the Polycomb group, and Trithorax proteins (Allis and Jenuwein, 2001 and Lacoste *et al.*, 2002). The three classes of enzymes that have been focused on in the last decade has been the histone acetyltransferases (HATs), histone deacetylases (HDACs) and the histone methyltransferases (HMT's).

University of Cape Town

Histone	Type of modification	Position	Residue(s)
H2A	Acetylation	NH_2- terminal tail	K5 and K9
	Ubiquitination	COO^- terminal tail	K119
H2B	Acetylation	NH_2- terminal tail	K5, K12, K15 and K20
	Ubiquitination	COO^- terminal tail	K123
H3	Acetylation	NH_2- terminal tail	K9,K14,K18,K23 and K 27
	Methylation	NH_2- terminal tail	R2,K4,K9,K17,K26 and K27
		Within loop1 between helix 1 and helix 2	K36
		Within the H3 globular domain	K 79
H4	Phosphorylation	NH_2- terminal tail	S10 and S28
	Acetylation	NH_2- terminal tail	S1 ,K5,K8,K12 and K16
	Methylation	NH_2- terminal tail	R20

Table 1: The covalent post –translational modifications of the core histones.

This table is representative of the known histone post-translational modifications. Many more could still be discovered. Most studied covalent modifications occur on the NH_2-terminal tails of histones except ubiquitination which occurred on the C-terminal carboxyl group. Certain lysine modifications also occurred within or near the histone H3 fold domain (Zhang *et al.*, 2003).

University of Cape Town

1.5.1 Histone Methylation

Histones have been known to be methylated for almost four decades (Van Holde 1989). These covalent modifications were known to occur on the ε- amino group of the lysine residues, and on the guanidino nitrogens of arginine residues (Table 1). The lysine residues on histones can either be mono-, di- or tri-methylated, whereas the arginine residues are only mono- or di-methylated (Zhang and Reinberg, 2001; Santa-Rosa *et al.*, 2002). These modification states on the lysine residues have no effect on the overall charge of the histone tail, but they do increase its basicity and hydrophobocity (Rice and Allis, 2001). It was thought that increasing the methyl groups from the mono- to tri-methylated state, there would be a higher affinity for DNA (Rice and Allis, 2001). It was shown that lysine methylation was dependent on two kinds of S-adenosyl-L-methionine (Ado Met) dependent enzymes, which were the *disruptor of telomeric silencing* (Dot1) and the SET methyltransferases (Xiao *et al.*, 2003). In addition arginine methylations were catalyzed by the PRMT/CARM methyltransferases (Zhang and Reinberg, 2001).

University of Cape Town

1.5.1.1 Arginine Methyltransferases

The covalent modification of arginine residues by methyltransferases resulted in either mono-or di-methylation. Di-methylated arginines can appear in symmetric or asymmetric configurations, and the enzymes identified in this process have been classified based on these two configurations. The first class, known as the protein arginine N^G-monomethylarginine, catalyzed the formation of the asymmetric N^G,N^G-dimethylarginine residues. The second class catalyzes the formation of

N^G-monomethylarginine and symmetric N^G, N^G-dimethylarginine residues (Gary and Clarke, 1998). The first class of arginine methyltransferases consisted of the protein arginine methyltransferase (PRMT1), PRMT2, PRMT3 and the co-activator-associated arginine methyltransferase 1 (CARM1/PRMT4).

Of these four arginine methyltransferases only PRMT1 and CARM1/PRMT4 have been identified as histone arginine methyltransferases, although they also displayed methyltransferase activity towards other substrates (Zhang and Reinberg, 2001).

PRMT1 has histone methyltransferase activity directed towards H4R3, and methylated RNA binding proteins in addition to the heterogeneous nuclear protein A1 (hnRNP A1) (Lee and Bedford, 2002).

The CARM1/PRMT4 complex showed histone methyltransferase activity towards H3 *in vitro* at R2, R17 and R26, and was a co- activator, which regulated transcription when co-recruited with the p160 family of co- activators. PRMT2 was identified *in silico*, by using the conserved catalytic core domain of previously identified arginine methyltransferases (Zhang and Reinberg, 2001). Recombinant PRMT3 was shown to have methyltransferase activity, but is yet to be fully characterized. The only class two enzyme identified thus far was PRMT5, which was shown to methylate histones H2A and H4, as well as the myelin basic protein and fibrillamin (Zhang and Reinberg. 2001). The arginine 3 methyltransferases appear to have a broad substrate specificity which regulates numerous cellular processes in comparison to the histone lysine methyltransferases (see below), which appear to have histones as their only substrate.

University of Cape Town

1.5.1.2 Lysine methyltransferases.

The first isolated mammalian SET methyltransferases were the human (SUV39H1) and the murine (SUV39h1) orthologues of the *Drosophila* position-effect variegation modifying proteins Su (var) 3-9 and the *Schizosaccharomyces pombe* Clr4 (Rea *et al.*, 2000). This group of proteins was characterized by their two conserved domains, which were the 60 amino acid chromo domain, and the 130-140 amino acid SET domain (Nislow *et al.*, 1997 and Rea *et al.*, 2000). The SET domain was found in proteins from various organisms ranging from viruses to human (Kouzarides, 2002 and Manzur *et al*, 2003). This evolutionary conserved domain was first characterized as a common structural motif in the PEV proteins *Su (var) 3-9*, the PcG protein *E (z)* and the trxG proteins *TRX* (Rea et al., 2000). The experiments by Rea *et al.*, (2000) showed that the catalytic activity of the SET domain was dependent on the adjacent cysteine-rich regions. The exception to this was the 119 amino acid SET domain protein from the large double stranded DNA-containing virus, *Paramecium bursaria chlorella,* which lacked these cysteine -rich pre-SET and post-SET motifs (Manzur *et al.*, 2003).

In recent reviews it was outlined that the SET domain was found in more than 300 proteins in a wide range of species. To date 6 Set domain proteins have been identified in *S. cerevisiae*, 11 in *S. pombe*, 41 in *D. melanogaster,* and there are 73 in the human database. (Kouzarides, 2002; Sims III *et al.*, 2003) In the review by (Kouzarides, 2002) the SET proteins were classed into four families based on their identity with the human SET domain and their relationship with the *S.cerevisiae* SET domain. The four families are the SET1, SET2, SUV39 and the Riz family. The dendogram shown in Figure 5 was based on the similarity of SET proteins to the yeast Set1p methyltransferase.

University of Cape Town

1.5.2 The SET protein families

1.5.2.1 The SUV39 family

The SUV39 family is characterized in having both pre-SET and post-SET cysteine - rich domains (Kouzarides, 2002). As previously mentioned, the human SUV protein SUV39H1, and the murine SUV protein, Suv39h1, were the first to be isolated and characterized as histone lysine methyltransferases (Rea *et al.*, 2000). These proteins are known to methylate H3 at lysine 9 and 27 (Sims III *et al.*, 2003).The methylation of H3K9 was shown to result in the binding of the heterochromatin protein (HP1) through its chromo domain, which was involved in transcriptional silencing at euchromatin, and maintenance of heterochromatin (Bannister *et al.*, 2001; Nakayama *et al.*, 2001; Nielsen *et al.*, 2001).

University of Cape Town

1.5.2.2 SET1 family

The Set1 family consists of all the proteins represented in Fig 6 and unlike the SUV39 family, do not contain a pre-SET cysteine - rich domain (Kouzarides, 2002). The Set1p protein from *S. cerevisiae* was shown to be a H3K4 methyltransferase (Roguev *et al.*, 2001). The Set 1 family can be divided into two sub- families, based on function and homology to the PcG and Trx groups of proteins. The sub family SET1A is representative of the Trx group. It includes the human mixed lineage leukemia (MLL), *Drosophila* trithorax (TRX), and human trithorax, and have been shown to methylate H3K4 and was associated with transcriptional activity (Sim III *et al.*, 2003). The SET1B represents a subfamily of PcG.

1.5.2.3 The SET2 family

The yeast SET2 and the mammalian nuclear receptor binding SET-domain protein (NSD1) have both been shown to methylate H3K36. They have structural and functional similarities to both the SET1 and SUV39 families. The *Drosophila absent, small or homeotic discs* (ash1) gene is 32% identical to the SU (VAR) 3-9 proteins in having both a pre-SET and post-SET domains (Byrd and Shearn, 2003). It was shown by Beisel *et al.*, 2002 that the Ash1 protein had histone methyltransferase activity *in vitro* to H3K9 and H4K20.

1.5.2.4 The Riz family

This family consists of the retinoblastoma-interacting zinc finger (RIZ), the B-cell induced maturation protein (BLIMP1) and the PFM1 proteins. Unlike the aforementioned SET families, the RIZ family has the SET domain situated near the C-termini, and there are no pre-SET and post-SET cysteine-rich domains. In addition, members have insertions within the NHSC catalytic motif, which would prevent any catalysis. The RIZ proteins have a zinc finger motif, which implies that they may have DNA binding activity, and they may function as transcription factors (Kouzarides, 2002).

It appears that the classification of the different SET proteins within the SET1, SET2 and SUV39 families is premature. As mentioned the EZ, EZH2 and ASH1 proteins are placed into families based on sequence similarities to humans, and display functional characteristics of proteins within other families. In future, as more data becomes available on these proteins, a classification should be based on the various domains and the location of the SET domain within these proteins.

University of Cape Town

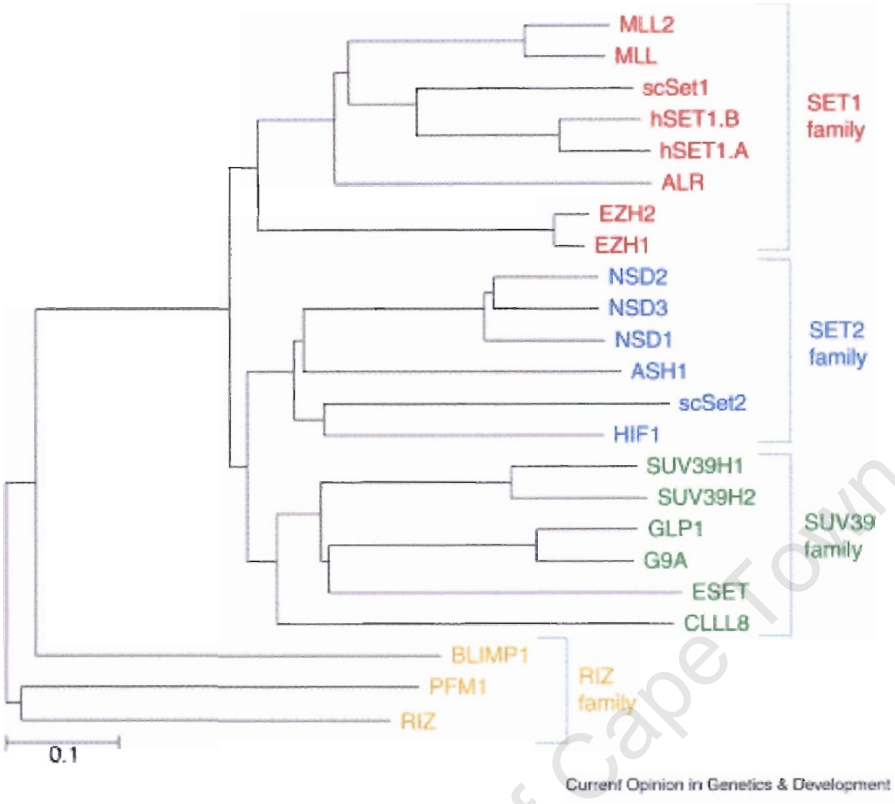


Figure 5: The Set family of transcriptional activators and repressors.

A dendogram depicting the relationship between the SET family members. The non-rooted tree was based on a multiple sequence alignment of the SET domain from *S. cerevisiae* Set1p protein. The SET proteins can be grouped into four classes based on their homology within their SET domain (Kouzarides, 2002).

University of Cape Town

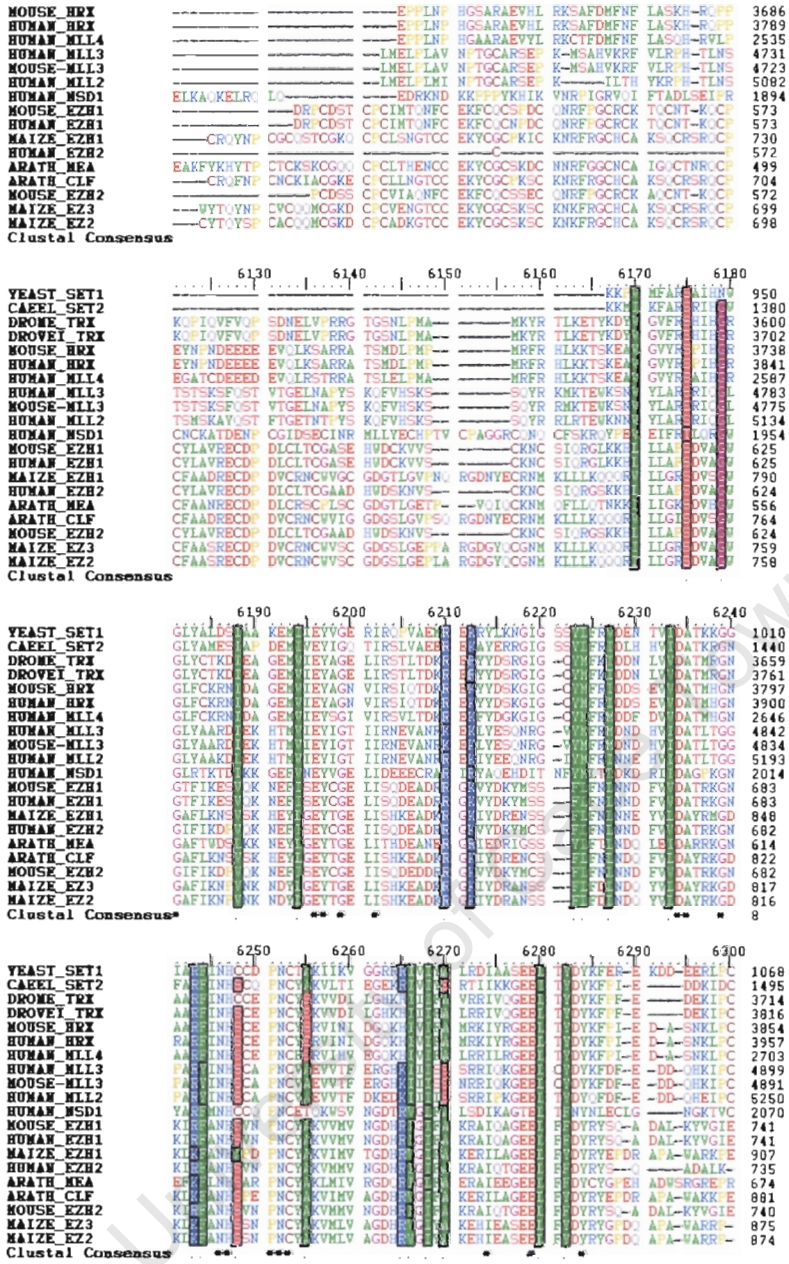


Figure 6: A multiple sequence alignment of the Set1 family

A multiple sequence alignment of the Set1 family showing the highly conserved Set domain. The clustal W program was used to generate a multiple sequence alignments. The black asterisks represent the conserved residues of the set1 methyltransferase catalytic core. Residues that are identical or have similar properties are shaded.

1.5.3 SET methyltransferase structure

During the past two years many advances have been made in obtaining 3D structures of the SET histone lysine methyltransferase domain. Crystal structures have been elucidated for the yeast *Sz. Pombe* Clr4, human SET7/9, *Neurospora crassa* DIM-5, *Paramecium bursaria* vSET and the non-Set containing methyltransferase Dot1p (Fig 7) (Wilson *et al.*, 2002) *al.*, 2002; Zhang *et al.*, 2002; Min *et al.*, 2002; Manzur *et al.*, 2003). One of the first crystal structures obtained was the yeast Clr4 methyltransferase, which is a member of the SUV39 group/family (Kouzarides, 2002). This group was shown to methylate H3K9 which was involved in gene silencing through the binding of the HP1 protein. The crystal structure of the yeast Clr4 methyltransferase at 2.3-3.0 Å resolution provided the first insights into the function of this protein, and how the SET domains facilitated the catalytic process (Min *et al.*, 2002).

University of Cape Town

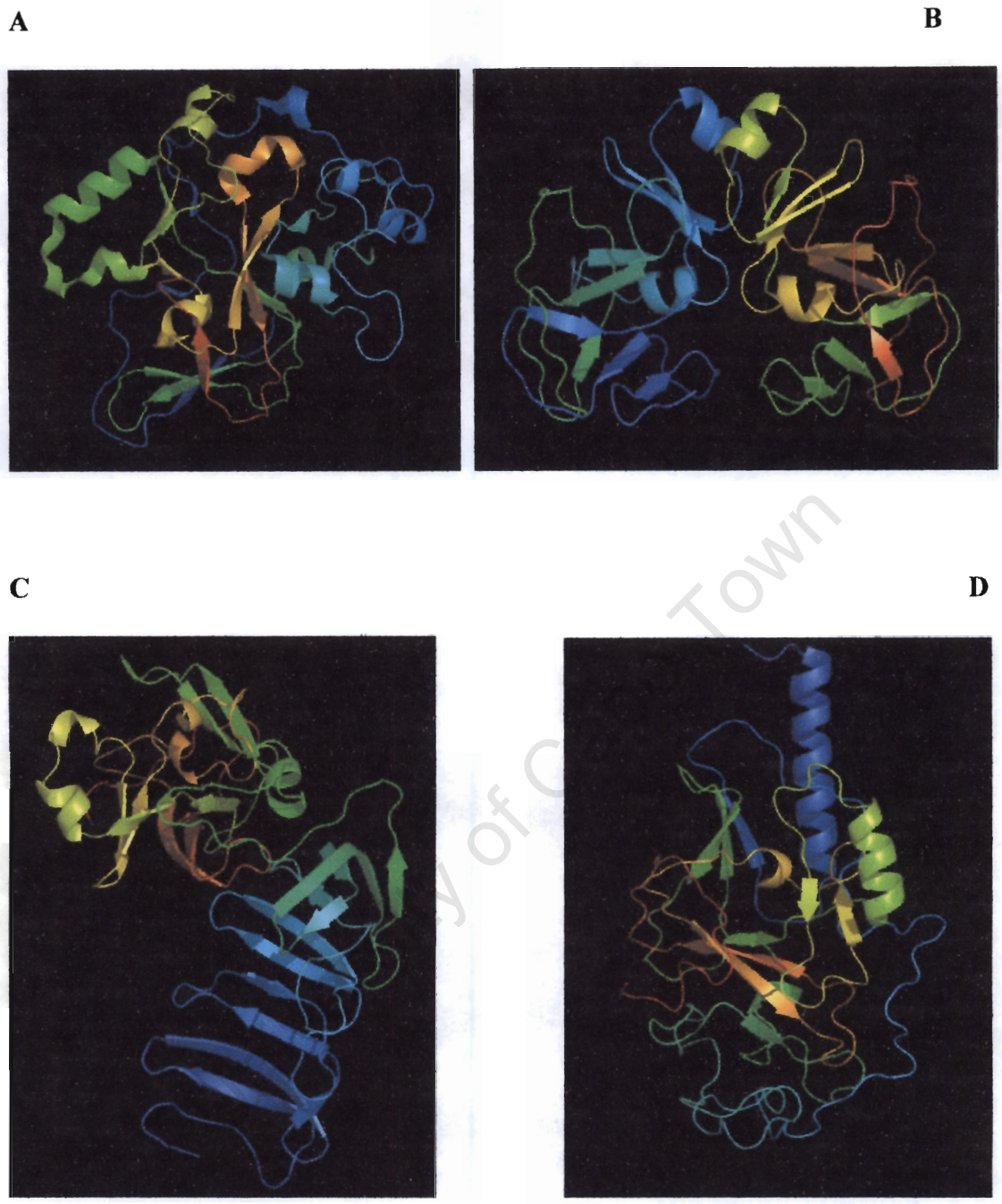


Figure 7: The crystal structures of four histone lysine methyltransferases

(A) *Neurospora crassa* DIM-5, **(B)** *Paramecium bursaria* vSET **(C)** human SET7/9 **(D)** yeast *Sz. Pombe* Clr4. These methyltransferases have structural similarities within their C-terminis, contains the evolutionary conserved SET domain (Zhang *et al.,* 2003; Manzur *et al.,* 2003, Min *et al.,* 2002 and Kwon *et al.,* 2003).

1.5.3.1 The structure of *the Sz. pombe* Clr4 methyltransferase

The crystal structure by Min *et al.*, (2002) allowed several deductions. The Clr4 enzyme consisted of an N-terminal (pre-SET) and a C-terminal (post-SET) domain, which flanked a 192-490bp catalytic SET domain (Fig 7D). The N-terminal region that preceded the pre-SET domain, consisted of one helix (αA), one β-strand (β1) and a long loop (loop2). The pre-SET region consisted largely of unstructured random coils, and also included a short helix (αB) and a short β-strand (β2). The post –SET domain consisted of a 30 residue unstructured region (Fig 8B). The SET domain consisted of eight β-sheets (β3-β10) and two short helices (αC and αD) (Fig 8A&B). The catalytic site of the SET domain was negatively charged, and was situated in a shallow groove. The residues 406-412 (Arg-Phe-Phe-Asn-His) formed the left side of the shallow groove situated in the loop connecting helix αD with β9. The Clr4 proteins predicted that S-adenosyl-L-methionine's binding pocket (Fig 8A green arrow) was not situated in a solvent- devoid region as in the S-adenosyl-L-methionine (SAM) dependent DNA and protein arginine methyltransferases. Based on biochemical evidence, the SUV39/Clr4 family has sequence conservation at residues 319-325, which is important for the stabilization of a zinc cluster and residues in the pre-Set region (Fig 8B black arrow) (Min *et al.*, 2002). Thus far the only structure which demonstrated the catalytic mechanism was the SET7/9 methyltransferase, which was solved with a bound substrate molecule (Kwon *et al.*, 2003).

University of Cape Town

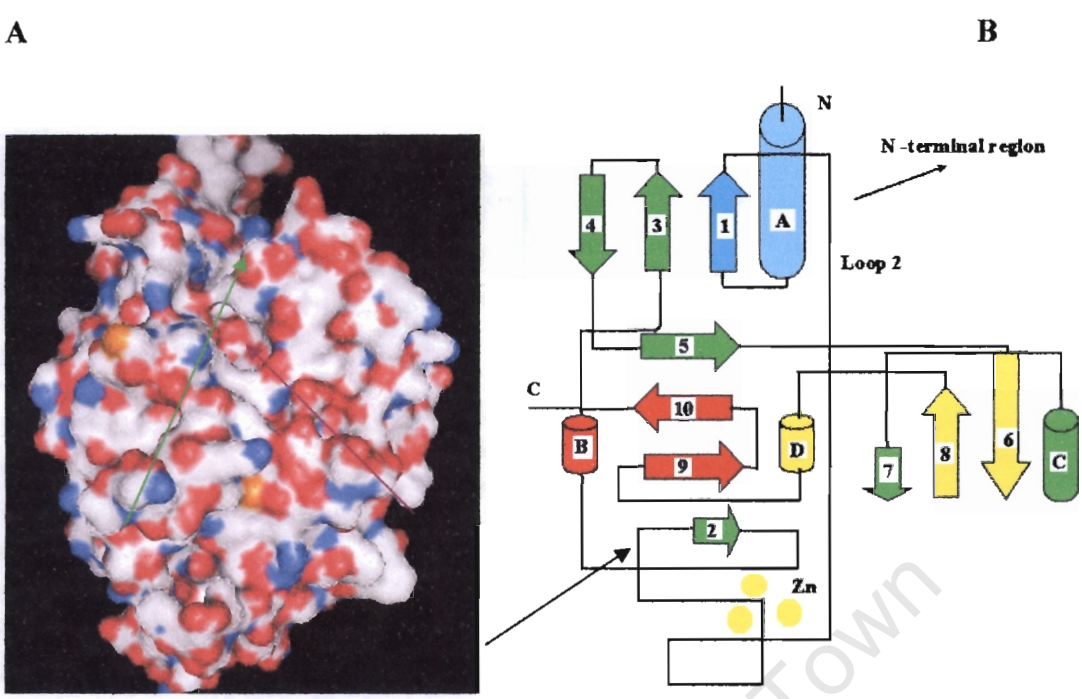


Figure 8: The structure of the Clr4 methyltransferase

(A) A surface representation of the of the Clr4 structure where the arrows indicates the predicted catalytic site. **(B)** Schematic organization of the SET domain which consisted of 10 β-sheets and 4 α-helices (Min *et al.*, 2002).

University of Cape Town

1.5.3.2 The catalytic mechanism of a SET methyltransferase

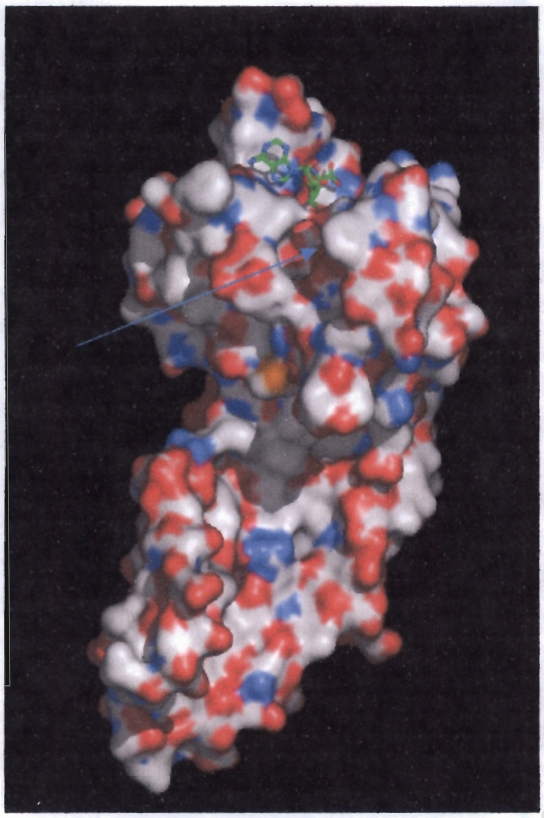
Due to the high sequence conservation among the SET family of proteins and their structural similarities, it is apparent that these characteristics will be evident in their catalytic sites. Thus far only the human SET7/9 –S-adenosyl-L-methionine mechanism of lysine methyl transfer has been revealed through a high resolution crystal structure (Fig 8A and B) (Xiao *et al.*, 2003; Kwon *et al.*, 2003; Wilson *et al.*, 2002). From these findings we can extrapolate that the histone binding site in Fig 8A (red and green arrows)

is where catalysis should take place. The human SET7/9 protein was shown to methylate histone H3 lys4 (H3K4). This protein can only mono- and di- methylate H3K4 and is present at both active and inactive genomic regions (Sims III *et al.*, 2003).

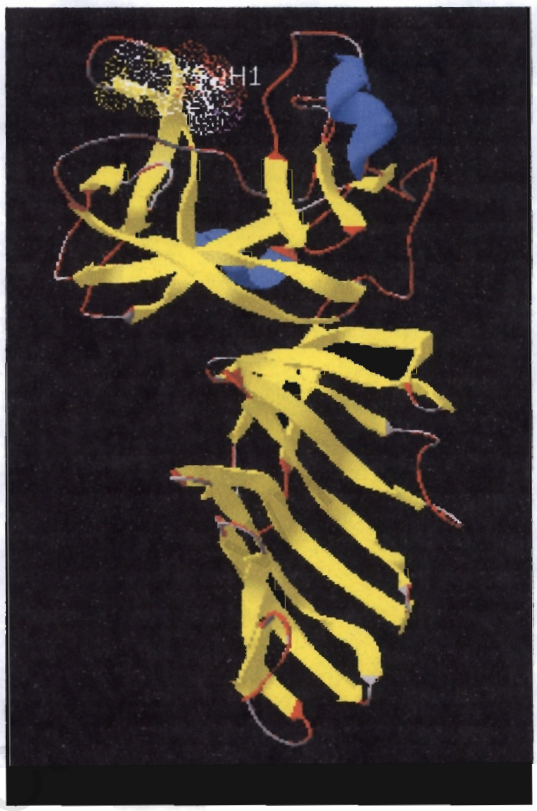
The enzyme – substrate data by Kwon *et al*, (2003) and Xiao *et al*, (2003) gives us an indication of the mechanism by which catalysis would take place. The Adomet binding site consisted of a channel with dimensions 9.5Å X13Å and a depth of 8.5Å formed by β-sheets and variable loops (Fig 9A & B). The channel surface has the R/HxxNHS motif, which is highly conserved among the SET domain family. The Adomet molecule interacted with this channel via hydrogen bonding and Van der Waals interactions. The binding of the histone tail occurs in a channel which has a diameter and a depth of 6Å. This channel is formed by the aromatic side chains of tyrosine residues (Y245, Y305, Y335 and Y337) as well as L267 (Fig 9A and C) (Kwon *et al.*, 2003 and Xiao et al 2003). Through site directed mutagenesis and ELISA Kwon *et al.*, (2003) showed that the residues H293, H297, K294, D276, W352, Y305, Y245 and Y335 are important for enzyme catalysis (Fig 9C and D) (Kwon *et al.*, 2003 and Xiao e*t al.*, 2003). Because of the structural similarities the yeast Clr4 methyltransferase probably catalyses H3 methylation in a similar manner to the human SET7/9.

University of Cape Town

A



B



University of C

C D

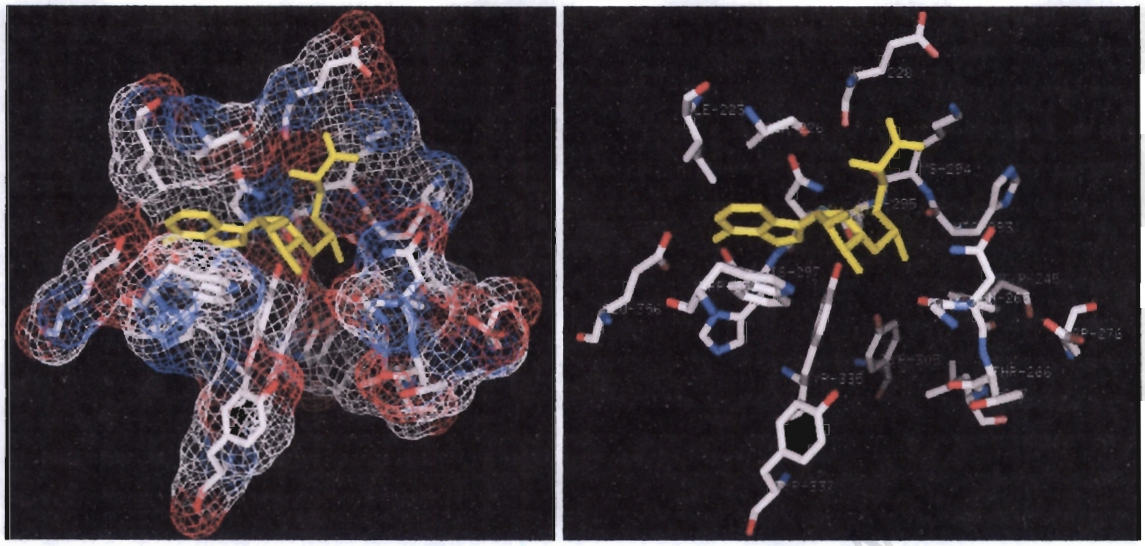


Figure 9: The catalytic site of a SET7/9 H3-K4 methyltransferase

(A and B) Surface and ribbon representation of SET7/9 binding an Adomet molecule and the histone tail binding site blue arrow **(C and D)** The SET7/9 catalytic site showing the interaction with a co-factor through van der Waals and hydrogen bonding (Kwon *et al.*,2003 and Xiao *et al.*,2003).

1.5.4 Functions of histone lysine methylation

Histone methylation has been linked to various functions in different organisms. The methylations on H3K9, H3K27 and H4K20 have been linked to gene repression, DNA methylation, heterochromatin formation, and X-inactivation (Lachner and Jenuwein, 2002). In contrast, the methylation of H3K4, H3K36 and H3K79 are associated with transcriptionally active regions (Sims III *et al.*, 2003).

Histone H3 lysine 4 methylation has been shown to regulate transcription through its association with the Iswi ATPase and is involved in rDNA and telomeric silencing (Briggs *et al.*, 2001; Bryk *et al.*, 2002; Santos Rosa *et al.*, 2003). The regulation of these

University of Cape Town

transcriptional events and the silencing of the above mentioned regions are not dependent on the silencing in regulation (SIR) proteins. Hui Ng *et al.*, (2003) has shown that the regulation of transcriptional events is dependent on the level of H3-K4 methylation at the 5' ends of genes. The methylation of coding regions could prevent the spreading of heterochromatin and could provide an epigenetic memory for past transcriptional events (Bryk *et al.*, 2002; Ng *et al.,* 2003). Histone H3 methylation (K9 and K27) was clearly shown to be involved in gene silencing through the binding of the HP1 protein or through DNA methylation. Certain enzymes such as the human Ezh2 and the G9a methyltransferase have the ability to methylate H3K9 and K27 (Sims III *et al.*, 2003 and Cao *et al.,* 2002). It was shown in the fission yeast and higher eukaryotes that K9 methylation plays an important role in heterochromatin formation and gene silencing.

Through the characterization of various methyltransferases from different organisms, the role of histone lysine methylation was shown to function by maintaining transcriptional boundaries, and the regulation of gene expression (Table2). The field of histone methylation is growing rapidly since it has been shown that these modifications and the enzymes responsible for the catalysis are implicated in various pathologies.

University of Cape Town

Organism	Site Methyl-ated	Methyltransferase	Function	Reference
Yeast	H3K4	spSet1p	Telomere maintenance and DNA repair	Kanoh *et al.*,2003 Noma and Grewal.,2002
		ScSet1p	rDNA silencing	Bryk *et al.*,2002
			Telomeric silencing	Krogan *et al.*,2002
			Gene expression	Boa *et al.*,2003:Santos Rosa *et al.*,2002; Miller *et al.*,2001; Carvin and Kladde,2004
	H3K9	Clr4	Centromeric heterochromatin and gene silencing	Bannister *et al.*,2001 Nakayama *et al.*,2001
	H3K79		Transcription activation	Sim III *et al.*,2003
Neurospora crassa	H3K9	DIM5	DNA methylation	Tamaru and Selker,2001
Arabidopsis thaliana	H3K9	Kryptonite	DNA methylation	Jackson *et al*, 2002
Drosophila melanogaster	H3K4	Ash1	Gene expression	Byrd and Shearn,2003

University of Cape Town

	H3K9	Su(var)3-9	Heterochromatin gene silencing	Schotta *et al.*,2002 Eskeland *et al.*,2004
Human	H3K9 &K27	E(Z)	Homeotic gene silencing	Czermin et al.,2002
	H3K4	SET9	Transcriptional regulation	Nishioka *et al.*,2002
		MLL, ALL-1 SET7/9	Gene activation	Sim III *et al.*,2003
	H3K9& K27	G9a	Euchromatic gene regulation	Tachibana *et al.*,2001;Litt *et al.*,2001
	H3K9	Suv39H1	Pericentric heterochromatin formation	Lachner *et al.*,2001;Vaute *et al.*,2002 Bannister *et al.*,2001
			Gene silencing	Nielsen *et al.*,2001
		SETD1	Gene silencing	Schultz *et al.*,2002

Table 2: The various histone methyltransferases and their functions.

Known histone methyltransferases that occur in humans, yeasts, *Arabidopsis thaliana*, *Neurospora crassa* and *Drosophila melanogaster*.

University of Cape Town

1.6 The Histone code

Through the discovery of the various enzymes responsible for the covalent modification of histones, many advances have been made in understanding their role in the epigenetic regulation of genes. This led Jenuwein and Allis, (2001) to propose a "histone code hypothesis", suggesting that covalent modifications of the histone N-terminal tails could induce interaction affinities for various chromatin associated proteins. The experiments of Nielsen *et al.,* (2001) provided the first evidence for the histone code hypothesis. They showed that methylation of H3K9 by the *D. melanogaster* SUV39H1 methyltransferase created a binding site for the chromo-domain of HP1. These proteins, together with the retinoblastoma (Rb) proteins, repress the cyclin E promoter. In contrast, it has been shown that in *S.cerevisiae* the Set1 histone methyltransferase is responsible for rDNA silencing which is not mediated by the SIR2 mechanism (Bryk *et al.*, 2001).

Various combinations of histone post-translational modifications may occur on the same or different histones tails. The methylation of H3K4 and acetylation of H3K9 on the same N-terminal tail has been shown to be involved in transcriptional activation during developmental changes at the chicken β-globin locus (Litt *et al.*, 2001). Subsequently, phosphorylation of S10 and methylation of K9 of histone H3 was also shown to be involved in transcriptional activation (Nowak and Corces, 2004). The ubiquitination on H2B K123 was required for the methylation of H3K4 by the Set1 methyltransferase, and was required for the localization of this enzyme at the 5'ends of genes (Sun and Allis, 2002; Hui Ng *et al*; 2003; Daniel *et al.*, 2003).

University of Cape Town

The concentration and various combinations of these post-translational modifications were responsible for the formation of higher order chromatin structures such as euchromatin and heterochromatin. For example, histone lysine methylation was shown to be involved in heterochromatin formation in contrast to acetylation which is associated with euchromatin (Jenuwein and Allis. 2001).

University of Cape Town

1.7 Project aims

It was shown that the yeast *Saccharomyces cerevisiae* Set1p methyltransferase methylated histone H3 at lysine 4 (Boa *et al.*, 2003; Roguev *et al.*, 2001) It has been established that the absence of the Set1p methyltransferases caused repression of genes within yeast (Boa *et al.*, 2003). Methylation by the Set1p methyltransferase was responsible for 50% of the total H3K4 methylation and has been shown to occur at both inactive and active genes, with tri-methylation occurring only at the 5'end of active genes (Bernstein *et al.*, 2002; Santa Rosa *et al.*, 2002). It was not clear whether the repression of genes within the yeast genome was due to the global change in chromatin structure or due to the recruitment of various transcription factors.

In order to understand the role of K4 methylation, specifically whether it maintained a general, transcriptionally condusive structure by virtue of the chemical modification, or whether it allowed the recruitment of chromatin associated factors, we investigated the role of H3K4 methylation on bulk chromatin structure.

To analyze the chromatin structure in a wildtype and a *set1*$^{-}$ strain, nuclei and sphaeroplasts was isolated. These nuclei and sphaeroplasts were treated with different concentrations of micrococcal nuclease (Mnase) to determine the accessibility of this enzyme to chromatin. To determine the changes of the nucleosome structure at individual promoter regions, genes that showed a significant increase and decrease in transcriptional activity in response to deletion of *the SET 1* gene was analyzed. It was expected that these experiments would contribute to the understanding of the role of H3K4 methylation in transcription.

University of Cape Town

CHAPTER 2: MATERIALS AND METHODS

The following molecular and biochemical manipulations were performed using established protocols (Ausubel *et al.*, 1995; Adams *et al.*, 1997 and Kent *et al.*, 1993). All the chemical reagents and media were of molecular biological grade and were prepared using deionised water produced with a milli-Q filtration system (Milli-Pore). All synthetic DNA probes were synthesized by the Oligonucleotide Synthesis Unit in the Department of Molecular and Cell Biology, University of Cape Town. Plasmid and gel extraction purifications were performed using QIAquick plasmid and gel purification kits according to the manufacturers instructions (Qiagen, GmbH, Hilden, Germany).

University of Cape Town

2.1 Yeast strains

All nuclei and sphaeroplasts isolations were performed using the strains LPY 197 (MAT **a** *ade 2-101 his3Δ-200 lec12Δ1 trp1Δ1 lys2-801 TELadh4::ura3*) and the LPYΔ set1a, is isogenic except at the SET1 locus (*set1:: His3*) (Nislow *et al.*, 1997 and Boa *et al.*, 2003).

2.2 Cell growth

Yeast media was prepared as described previously (Adams *et al.*, 1997). The rich medium yeast peptone dextrose (YPD) contained 1% (w/v) yeast extract (Difco), 2% (w/v) peptone and 2% (w/v) glucose. The complete synthetic medium lacking histidine contained 0.67% (w/v) yeast nitrogen base without amino acids (Difco), 0.77g/L CSM-HIS, 2% glucose and 1% (w/v) yeast extract. All the above media

contained 0.001% (w/v) polypropylene glycol 2000 (Fluka) to minimize flocculation. For agar plates, 2% (w/v) bacto-agar was added and all the media autoclaved for 15 minutes. Yeast cells were inoculated into 5 ml YDP and grown at 30°C overnight with rotation at 150 rpm (Sanyo orbital shaker). These fresh overnight cultures were used in subsequent inoculations (Adams *et al.*, 1997).

2.3 Yeast nuclei Isolations

Nuclei preparations were performed according to Allis laboratory protocol with slight modifications. Micrococcal nuclease (MNase) digestions were performed as described by (Szent-Gyorgyi *et al.*, 1987).

Approximately 100µl of a pre-inoculum was taken, and inoculated into 2000ml of YDP and grown at 30°C to an OD 0.7-0.8 (approximately 16 hours). Cells were harvested by centrifugation at 5000rpms for 5min at 4°C (JA 14 rotor). The pellets were washed twice in 50ml of ice cold water and centrifuged. The cells were resuspended in 10ml of sphaeroplasting buffer (1M sorbitol, 50mM potassium phosphate (pH6.5), 14mM β-mercaptoethanol), and 1ml of 10mg/ml Zymolyase (Seikagaku American Inc.) added. Sphaeroplasting was allowed for 60min at 30°. To determine sphaeroplasting completeness, the ability to "squish" cells was determined with the aid of a light microscope. Sphaeroplasts were pelleted at 5000 rpm for 5 min at 4°C, resuspended in 10ml sphaeroplasting buffer, and centrifuged at 5000rpm for 5min at 4°C. Sphaeroplasts were then resuspended in 10ml of lysis buffer (18% Ficoll 400, 20mM potassium phosphate (pH 6.8), 1mM $MgCl_2$, 0.5mM EDTA, 1mM PMSF and 1µg/ml leupeptin/pepstatin) and homogenized with a Teflon-coated dounce homogenizer at

University of Cape Town

50rpm. The cell suspension was then chilled on ice for 10mins, and cell debris pelleted at 5000rpm for 10min at 4°C.The supernatant was transferred to a new centrifuge tube, and the nuclei pelleted at 18000rpm for 30mins at 4°C. The nuclei were resuspended in 10ml NP buffer (0.34M sucrose, 20mM TRIS (pH 7.4), 50mM KCl, 5mM $MgCl_2$ 1mM PMSF and lμg/ml leupeptin/pepstatin) and pelleted in 2ml cushion buffer (NP buffer containing 1.7M sucrose) at 15000rpm for 30min at 4°C. Nuclei were finally resuspended in 1ml NP buffer containing 50 %(v/v) glycerol and stored at -20°C until use. Nuclei that were to be used directly for micrococcal nuclease digestions were resuspended in 4ml of digestion buffer D (10mM Hepes (pH 7.5), 0.5mM $MgCl_2$, 0.05mM $CaCl_2$ and 1mM PMSF).

2.3.1 Nuclei Quantitation

Nuclei were quantitated by diluting the nuclei 1:100 in 0.1M NaOH. Spectrophotometer readings were taken at 260nm, where 1mg nuclei = 20 OD unitsA_{260}. Concentrations for the wildtype and mutant strains were equilibrated with buffer D.

University of Cape Town

2.3.1 Micrococcal nuclease digestions of Nuclei

The resuspended nuclei were divided into 200µl aliquots and 10µl of micrococcal nuclease (Mnase) was added using the concentrations shown in Table 3.

Sample	Stock (20U/µl) dilution	[Micrococcal nuclease] (U/ml)	[Final] (units)
1	15	15000	750
2	7.5	7500	375
3	3.75	3750	188
4	1.88	1880	94
5	0.94	940	47
6	0.49	490	25

Table 3: The treatment of yeast nuclei with different concentrations micrococcal nuclease.

Micrococcal nuclease digests were done at 37°C for 15 min and the digests were quenched in 1/10 volume 100mM EDTA. The samples were then adjusted to 20µl of 1mg/ml Ribonuclease H (Sigma) and incubated at 37°C for 30minutes. The samples were quenched by adding 10µl of 5M NaCl, 30µl of 20% SDS and 30µl of 10mg/ml of Proteinase K (Sigma). The mixture was incubated at 37°C for 3hrs.

University of Cape Town

2.3.2 DNA purification

DNA extraction and purification was performed using standard phenol extraction and ethanol procedures (Ausubel *et al.*, 1995). The DNA was resuspended in 20µl of dH_2O.

2.4 Micrococcal nuclease digestion of permeabilised yeast sphaeroplasts

A 100ml volume of YPD was inoculated with 100µl of fresh over-night culture and incubated for 16hrs (OD_{600} ~ 0.6-0.8) at 30° C with shaking at 70 rpm (Sanyo orbital shaker). The yeast cells were harvested by centrifugation at 5000rpm for 5 min at 4°C. Cells were resuspended in 20ml of 1M sorbitol, and at 5000 rpm for 5 min at 4°C. Cells were then resuspended in 10ml of sphaeroplasting buffer (1ml of 10mg/ml zymolyase, 10mM β-mercaptoethanol and 1M sorbitol) and incubated at 30°C for 1hr. The yeast sphaeroplasts were harvested by centrifugation at 5000 rpm for 5min at 4°C. The sphaeroplasts were washed in 1M sorbitol and re-pelleted. The sphaeroplasts were resuspended in 800µl of Buffer A (1M sorbitol, 50mM NaCl, 10mM Tris-Cl (pH 7.4), 5mM $MgCl_2$, 1mM $CaCl_2$, 1mM β-mercaptoethanol and 0.5mM spermidine). Aliquots of 100µl of resuspended sample were dispensed into 6 eppendorfs and each sample mixed with 100µl of ice-cold Buffer B (Buffer A + 10 % (v/v) NP40). For the free DNA control an aliquot of 400µl of sphaeroplasts solution were dispensed into an eppendorf and mixed with 400µl ice cold Buffer B.

From the following micrococcal nuclease stocks, 10µl were used in micrococcal nuclease digest. The following concentrations of Mnase were used: 15U/µl, 7.5U/µl, 3.75U/µl, 1.88U/µl, 0.94U/µl and 0.49U/µl. The digests were incubated at 37°C for exactly 10min and quenched by adding 20µl 250mM EDTA (pH 8.0)/5 %(w/v) SDS followed by the

University of Cape Town

addition of 6µl proteinase K (10mg/ml). This mixture was further incubated at 37°C for 1hr. Standard phenol/chloroform extraction followed by ethanol precipitation was done as described in section 2.3.2. To the undigested DNA sample, 400µl of Buffer A was added and 100µl dispensed into 4 eppendorfs, to which 100µl Buffer B was added. For Mnase digests 2µl of Mnase was added from the concentrations listed in Table 4 (Kent *et al.*, 1993).

Sample	**Stock (10U/µl) dilution**		**[Mnase] (U/ml)**		**Final (U/ml)**	
	A	**B**	**A**	**B**	**A**	**B**
1	2.0	10.0	4000	20000	20	100
2	0.75	2.0	1500	4000	7.5	20
3	0.20	1.0	400	2000	2.0	10
4	0.10	0.2	200	400	1.0	2.0

University of Cape Town

Table 4: Genomic DNA was digested with the following concentrations of Mnase.

The concentrations of Mnase used in genomic DNA digests and for indirect end-labeling. The Mnase concentration in column B was used in experiments analyzing chromatin structure at the *TRX2* promoter using the 513bp probe. The Mnase concentration in column A was used in experiments for analyzing both the *TRX2 and SIR4* promoter region for indirect end labeling.

Incubate for 5-10 min at 37°C. Stop the reaction by adding an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1). DNA was then isolated using standard Phenol/Chloroform and ethanol precipitation method as described in section 2.3.2.The DNA was resuspended in 10µl of dH_2O.

2.5 Agarose gel electrophoresis

All micrococcal nuclease digest, naked DNA digests and PCR products were confirmed by agarose gel electrophoresis on a mini 1% (w/v) agarose gel using TBE electrophoresis buffer (90mM TRIS, 89mM Boric acid and 2.5mM EDTA) containing 10µg/ml of ethidium bromide. DNA was electrophoresed at 90V for 1hr and visualized by short wavelength (254nm) UV transillumination. Large 1 % (w/v) agarose (20 X 25cm) TBE gels were run at 25V overnight (Ausubel *et al.*, 1995).

2.6 Restriction endonuclease digest

The following marker and subsequent chromatin and naked DNA secondary digestions were performed using (500 units) of the enzymes at (10U/µl) BstEII and (1000 units) of (10U/µl) DraI (Roche). The restriction digests of the marker were performed in 20µl volumes that contained 10µl of DNA (Section 2.4), 1.7µl 10X restriction enzyme buffer, 4.3µl H_2O and 10U/µl restriction enzyme. The reaction was incubated at 37°C for 3hrs. The restriction enzyme digests were verified by running 2µl on a 1 % (w/v) agarose gel (Kent *et al.*, 1993).

University of Cape Town

2.7 Yeast genomic DNA isolation for PCR

LPY197 yeast cells were grown overnight at 30°C in 5ml of YPD. The cells were pelleted at 5000 rpm for 5 min at 4°C and resuspended in 500µl of 1M sorbitol, 100mM EDTA (pH7.5). Sphaeroplasting of the cell pellets was achieved by adding 20µl of a 2.5mg/ml of zymolyase (Seikagaku American Inc.) and incubating for 1hr at 30°C. The sphaeroplasts were harvested in a bench top micro centrifuge at 12000rpm for 1min. Sphaeroplasts were then resuspended in 500µl of 50mM Tris-Cl (pH7.4), 20mM EDTA and 500 µl 10% (w/v) SDS added, and incubated for 30min at 65°C. After the incubation, 200µl of 5M potassium acetate was added, and the mixture placed on ice for 1hr. Centrifugation was performed in a bench top micro centrifuge at 12000rpm for 5min. The supernatant was transferred to a new 1.5ml eppendorf and 1 volume of absolute isopropanol added. The mixture was inverted ten times, and allowed to incubate at room temperature for 5 min. Centrifugation was performed in a bench top micro centrifuge at 12000rpm for 10s. The supernatant was decanted, and dried in a speedivac for 45 minutes. The pellet was resuspended in 300µl of TE (10mM Tris, 1mM EDTA, pH 7.4), and 15µl of a 1mg/ml solution of Ribonuclease H added, and incubated at 37°C for 30min. An aliquot of 300µl of 3M sodium acetate (pH5.2) was then added, vortexed for 10s, and precipitated with 200µl of isopropanol. The DNA was pelleted by centrifugation at 12000rpm for 5min. The supernatant was discarded, and the pellet dried and resuspended in 100µl-300µl of TE (pH 7.4) (Adams *et al.*, 1997).

University of Cape Town

2.7.1 DNA quantitation

Purified DNA was spectrophotometrically quantitated by diluting the DNA in H_2O (1:10, 1:100, and 1:1000) and taking absorbance readings at 260nm in a spectrophotometer (Beckman). Concentration was calculated as in Ausubel *et al.*, (1995).

2.8 Polymerase Chain Reaction (PCR)

2.8.1 PCR Primers

Primers were designed for the *LMA1*/thioredoxin 2 (*TRX2*) and the *silencing in regulation* 4 (*SIR4*) gene promoter regions (Boa, *et al.*, 2003).

University of Cape Town

Primer name	Sequence position	Sequence
FTRX 2.P	**-339**	**5'-AAAGTC TTGTTGAGCCCGGTAA- 3'**
RTRX2.V3	**-170**	**5'-CGCTTAGTAAGAAAAGAGCC- 3'**
RTRX2	**+250**	**5'-CCTTACCGCCCTTGTAGAAGATTA- 3'**
FSIR 4.IF	**-280**	**5'-GCTATAACGTCCTTAAACATG- 3'**
RSIR4.V3	**-108**	**5' -CGCTTATTTCGTTATAGGTTT-3'**
RSIR4	**+233**	**5'-AACTGTTTATGGGAAGT CGTTTT -3'**

Table 5: Primer sequences used in the amplification of the *LMA1/TRX2* and *SIR4* promoter regions.

University of Cape Town

2.8.2 Amplification of TRX2 and SIR4 promoters

The amplification of the promoters for the *TRX2* and *SIR4* genes were performed by the Polymerase Chain Reaction (PCR) using 10-100ng of genomic DNA as a template in a 50µl reaction volume containing 5µl of 10X *pfu* DNA polymerase buffer (Promega), 1µl of 3U *pfu* DNA polymerase (Promega), 0.3µl of 150µM dNTP and 0.5µM of each primer. The thermo cycling parameters were as follows: an initial denaturation step at 94°C followed by 35 cycles consisting of 94°C incubation for 1min, annealing step of 50°C for 1min and an elongation step at 72°C for 2min. A final elongation step of 72°C for 5min

was performed and the reactions held at a temperature of 4°C until use. All PCR products were verified for correct size by agarose gel electrophoresis.

2.9 Indirect end labeling

2.9.1 Secondary digest of micrococcal nuclease digests

Restriction enzyme digests were performed by adding 7µl from a master mix containing 2µl H_20, 2.5µl 10X restriction enzyme buffers and 2.5µl of enzymes to each sample to be digested. The samples were incubated at 37°C for at least 3hr to overnight, and 5µl of 6X loading dye (Promega) was added to each of the digested samples. From each sample 2µl were loaded on a 1% agarose gel for confirmation of the completeness of the enzymatic digest. The remaining samples and 100bp molecular weight marker (Promega) were loaded on a large 1% agarose gel (20 X 25cm) run at 30V for 15hrs (Kent *et al.*, 1993)..

University of Cape Town

2.9.2 Southern blotting

The top right had corner of the agarose gel was clipped prior to denaturing the DNA within the gel. The gels were incubated in 10 gel volumes of denaturation solution (1.5M NaCl/0.5M NaOH) with gentle shaking at room temperature for two 15 min treatments. The gel was rinsed in 10 gel volumes of H_20 and the DNA neutralized by adding 10 volumes of neutralization solution (1.5M NaCl/0.5M Tris-Cl. (pH7.0)/1mM EDTA) (Kent *et al.*, 1993).

The blotting procedure was performed by upward capillary as follows:

The gel was transferred to a solid support consisting of 3X Whatman 3MM paper wick placed over a glass plate in a half filled dish with 20X SSC. The gel was placed on top of a wick and blotted against 20X SSC to a nylon hybridization membrane (GenescreenPlus, Amersham). A 15cm layer of paper towels was placed on top with a 0.2-0.4kg weighted object and left overnight. The transfer was confirmed by long wavelength UV transillumination. The DNA was UV cross-linked to the membrane and rinsed in 2X SSC at room temperature (Ausubel *et al.*, 1995).

2.9.3 Probe preparation

All probes were quantitated by a UV spectrophotometry and the DNA radiolabeled with 50µl [α-^{32} P] dCTP by the random prime method (Prime-A-Gene, Promega) using the accompanying protocol with one minor adjustment. The reaction was allowed to proceed for 3hrs instead of the recommended 1hr as this ensured a higher percentage of [α-^{32} P] dCTP incorporation. The unincorporated nucleotides were removed with a SigmaSpin Post-Reaction column using the accompanying protocol (Sigma). The percentage incorporation of the [α-^{32}P] dCTP was determined by liquid scintillation using a Beckman LS 5000 TD Liquid Scintillation Counter, using 1µl aliquots from before and after purification.

University of Cape Town

2.9.4 Hybridization of Probes

The nitrocellulose membrane was pre-hybridised in 10ml-20ml pre-hybridisation buffer (1.5X SSC, 5X Denhardt's, 0.1%SDS) at 64°C for 3hrs in a hybridization oven. The probe was boiled in 250µl 2.5mg/ml salmon sperm DNA for 5 min and the reaction

quenched on ice for 2 min. The probe/salmon sperm DNA mixture was added to pre-hybridisation buffer and hybridised at 64°C for 12-18 hrs in a hybridization oven.

The blot was washed once in 100ml wash buffer (2X SSC/0.1% SDS) for 20min at 60°C and twice in wash buffer at 64°C for 15 min. The blot was sealed in a transparent plastic bag and exposed to a small phosphor storage screen (Kodak) for 2 days. Visualization was performed using a Personal molecular imager FX Phosphoimager (Bio-Rad) and analysis performed using Quantity One software (Bio-Rad).

2.10 Reprobing of Southern blots.

Probed blots were placed in boiling 0.1% SDS and were then transferred to a hybridization oven for 1hr at 64°C with gentle agitation. The membrane was sealed in a transparent plastic bag and exposed to a small phosphor screen (Kodak) for 3 hrs. The clean membranes were then prehybridized and hybridized as above.

University of Cape Town

Chapter 3: Chromatin appears less susceptible to Mnase cleavage in the absence of histone H3 K4 methylation.

3.1 Introduction

The yeast *Saccharomyces cerevisiae* has become the preferred model organism for studying the association between histone post-translational modifications and transcription (Liang *et al.*, 2004). The yeast Set1p protein was identified as a histone methyltransferase specific for histone H3 lysine 4 (H3-K4) (Roguev *et al.*, 2001; Boa *et al.*, 2003). The microarray data from a *set⁻1* strain showed a decrease in gene expression in 80% of ORFs in the absence of the Set1p protein (Boa *et al.*, 2003). This correlates with the findings of Bernstein *et al.*, (2002) that showed that the di-methylation of H3K4 in coding regions relates with transcriptional activity. The experiments by Santos-Rosa *et al.*, (2002) indicated that tri-methylation occured only at the 5' end of active genes, while di-methylation occurs at both active and inactive genes. In contrast, the Set1 protein and H3K4 methylation maintained ribosomal DNA and telomeric silencing (Nislow *et al.*, 1997; Briggs *et al.*, 2001; Bryk *et al.*, 2002).

University of Cape Town

To address the question whether the main role of methylation of H3K4 by Set1p was to maintain chromatin structure in a specific state in *S.cerevisiae in vivo*, we isolated nuclei from wildtype and *set1⁻* strains and analyzed the chromatin structure nuclei by Mnase digestion. We proposed, as a first approximation, that Mnase susceptibility of the chromatin will reflect any major, stable, underlying structural changes in chromatin.

3.2 Results

3.2.1 Nuclei preparation

In order to determine the effect of H3K4 methylation on Mnase accessibility of chromatin we isolated nuclei from wildtype and *set1*⁻ strains, and digested the chromatin in each strain with different concentrations of Mnase. The nuclei concentrations were of equal concentration in both the wildtype and *set1*⁻ strains throughout the experiments and were equilibrated as explained in section 2.3.1. We were especially interested in apparent changes in digest extent at given concentrations, since this could represent differences in the accessibility and presumably structure of the underlying chromatin.

3.2.2 Micrococcal nuclease digestion: DNA structure and nucleosome arrangement

Micrococcal nuclease is an endonuclease which cleaves double stranded DNA in the absence of bound proteins. In these experiments the endonuclease cleaved the linker DNA between each successive nucleosome. This cleavage pattern would have the appearance of a DNA ladder when visualized on ethidium bromide stained agarose gel.

The gel in Figure 10 is a representation of the chromatin organization within a wildtype and a *set1*⁻ deleted strain. The nuclei were treated with increasing concentrations of micrococcal nuclease (MNase) (lanes 1-6), purified and run on a large agarose gel. All the variables in generating these data were kept constant for both the wildtype and the *set1*⁻ strain. From the ethidium bromide stained gel (Fig 10), it was clear that that the

University of Cape Town

nuclei were of a high quality, as both Mnase treated and naked undigested untreated genomic DNA had no appearance of degradation before digestion.

The presence of longer nucleosome oligomers at equal concentrations of Mnase in the wildtype strain suggested that the wildtype chromatin appeared to be more sensitive to MNase digestion than the *set1*⁻ deleted strain. In the wildtype there appears to be increased hypersensitivity to the MNase, which cleaves the DNA more rapidly and releases small mono and oligonucleosomes. By the analysis of each lane there appears to be uniform spacing beginning at lane1 to approximately lane3.

University of Cape Town

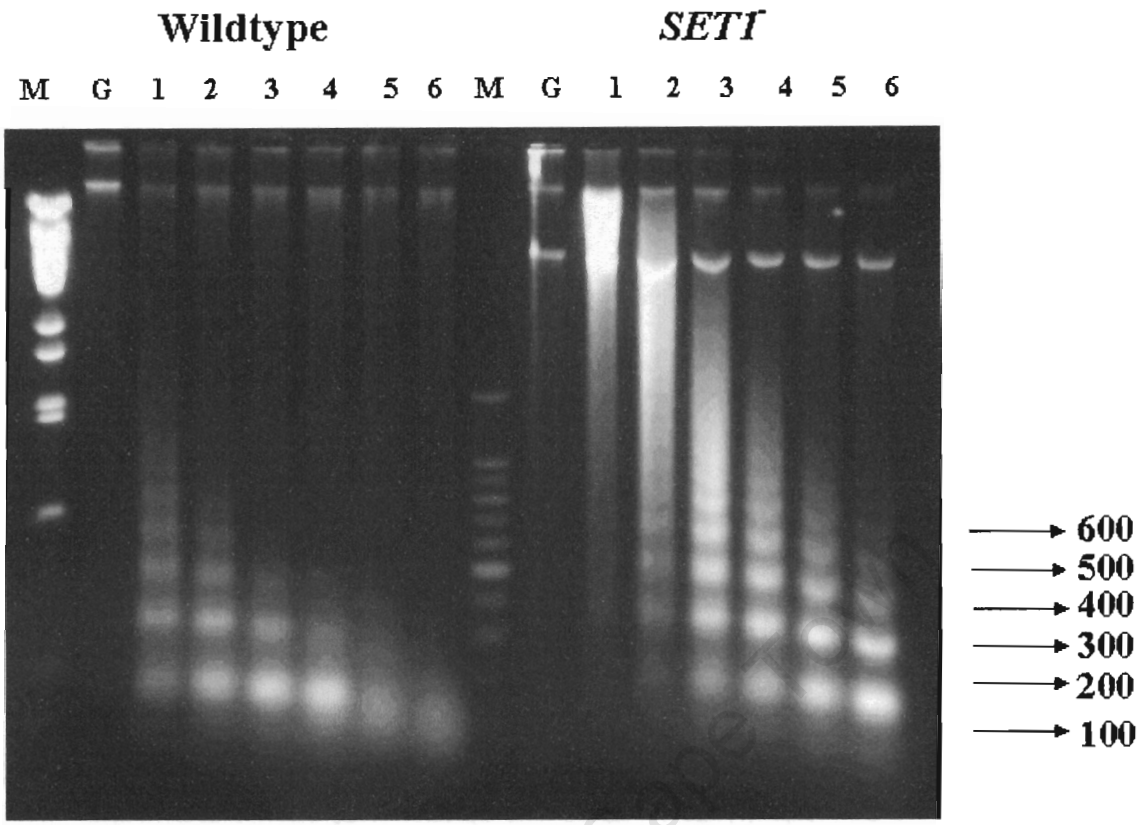


Fig 10: Chromatin structure in the wildtype and *set1*⁻ strains.

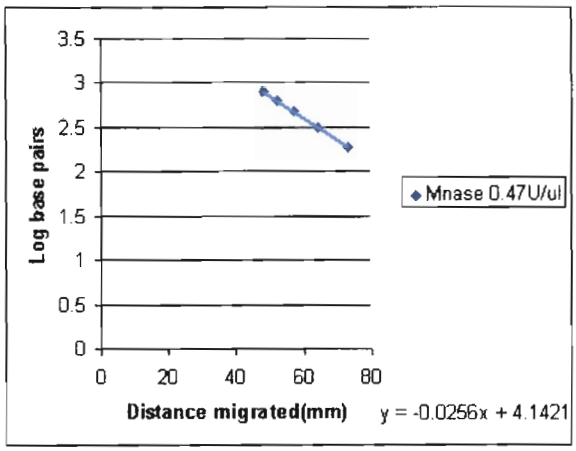
Chromatin in the wildtype and *set1*⁻ strains was digested with increasing concentrations of Mnase. (M) BstEII lambda marker, (G) genomic DNA Wildtype nuclei were digested with 0.47 U/µl (lane1), 0.94 U/µl (lane 2). 1.88 U/µl (lane 3), 3.88 U/µl (lane 4), 7.5 U/µl (lane 5), 15U/µl (lane 6) of Mnase. *Set1*⁻ nuclei were digested with 0.47 U/µl (lane1), 0.94 U/µl (lane 2) 1.88 U/µl (lane 3), 3.88 U/µl (lane 4), 7.5 U/µl (lane 5), 15U/µl (lane 6) Digestion of free DNA did not result in a regular pattern (data not shown). This is an ethidium bromide stained agarose gel visualized by UV trans illumination.

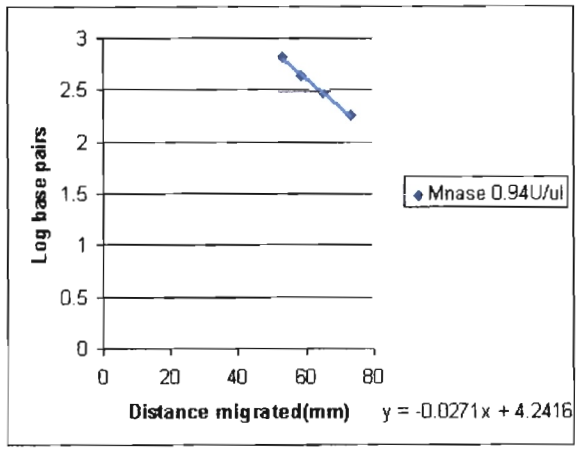
The *set1*⁻ strain appears to have uniformly spaced nucleosomes at the Mnase concentrations examined. The overall appearance of the DNA fragments show a sharp, intense banding pattern, with clearly observed mononucleosomes in comparison to the wildtype. Both the wildtype and *set1*⁻ strains appear to have the nucleosomes uniformly spaced, although wildtype begins to loose its uniformity with increased Mnase while the *set1*⁻ strain retains most of it at an equivalent concentration. If the Mnase concentration were increased in the *set1*⁻ strain, then we would have seen the "fuzzy" bands observed in wildtype lanes 4-6. The wildtype lane 2 has the same appearance as lane 6 in the set1⁻ strain.

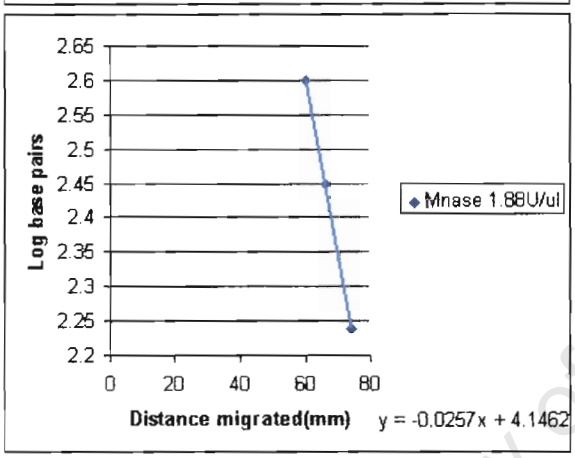
By observing the behavior on the graphs (Fig 11) for both the wildtype and *set1*⁻ strain, it therefore appears that the wildtype is more accessible to Mnase than the *set1*⁻ strain.

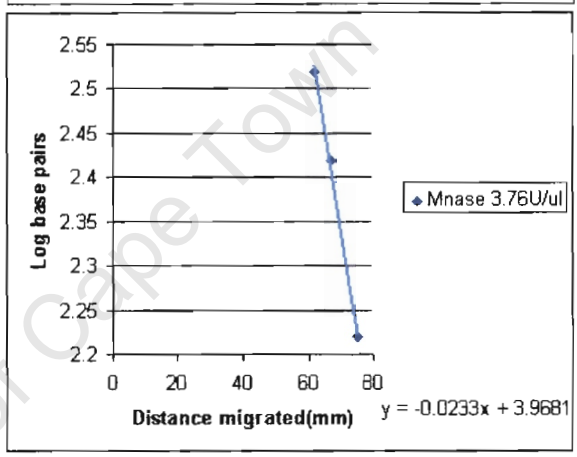
University of Cape Town

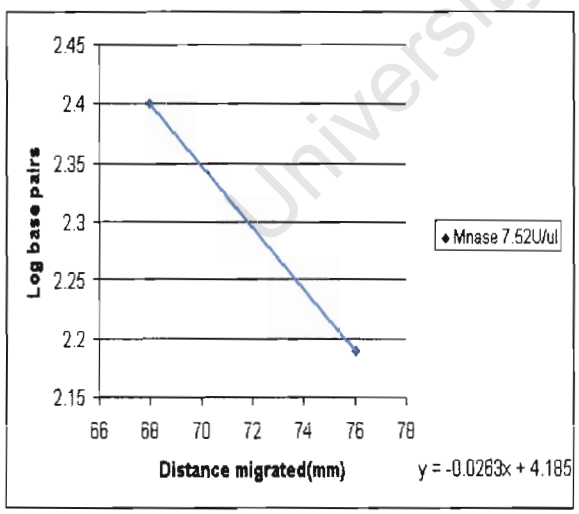
Wildtype





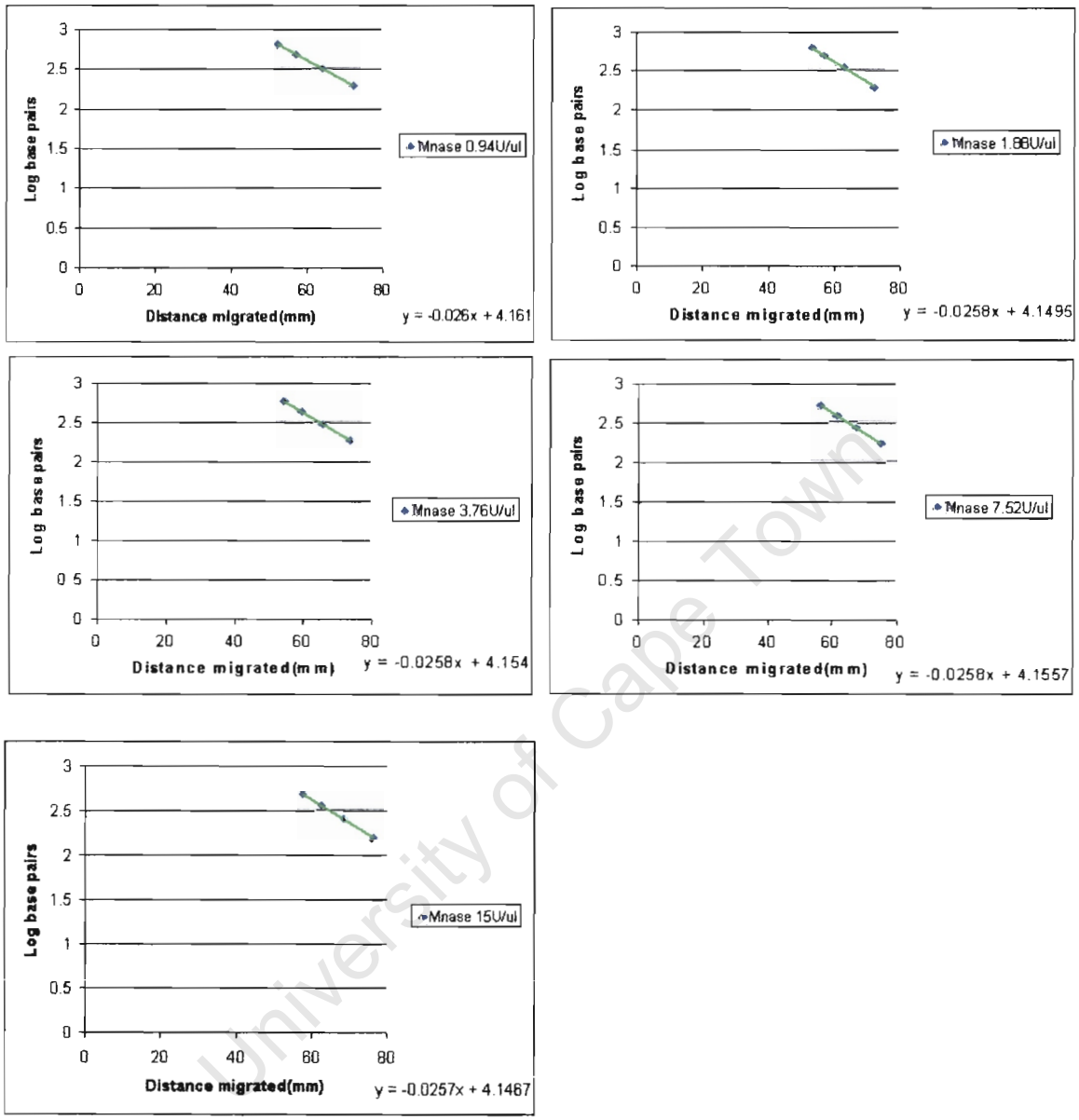






University of Cape Town

Set1^{-}



University of Cape Town

Figure 11: DNA oligomeric sizes as determined from 100bp molecular weight marker for a wildtype and set1^{-} strains

The data shown are sizes of DNA bands in (Fig10) for increasing concentrations of Mnase for nuclei isolated for a wildtype and set1^{-} strain. Digestions were done with yeast nuclei and sizes determined from an agarose gel.

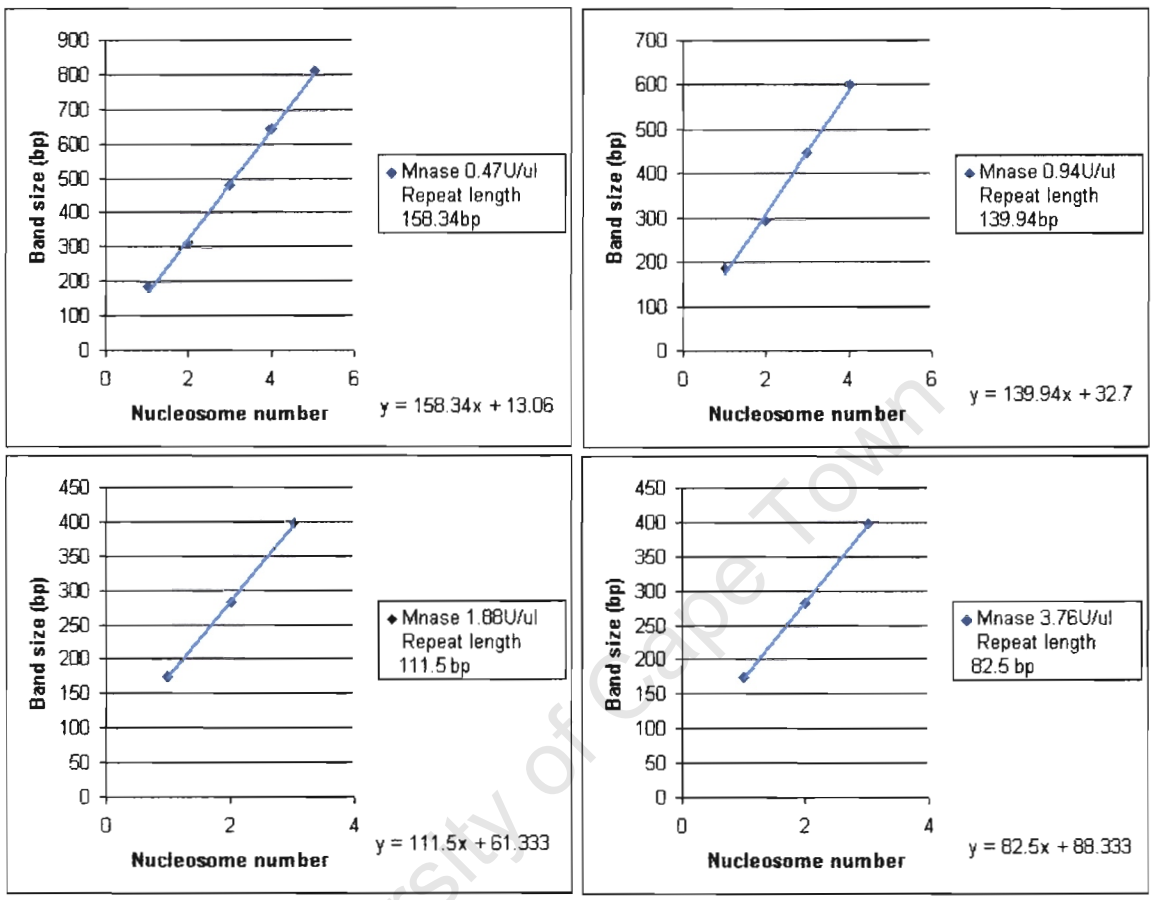
3.2.3 Nucleosome repeat length for the wild type and a *set1*⁻ strain

The nucleosome repeat length for the yeast *S.cerevisiae* during log phase growth is approximately 165 ± 5 (Hörtz and Zachau, 1980; Thomas and Furber, 1976). This represents the dyad to dyad spacing between nucleosomes i.e., the DNA associated with the nucleosome core (~146bp) along with the linker DNA (~20bp) on one side (Van Holde *et al*, 1979).

Figure 12 depicts the repeat length of DNA fragments digested with increasing amounts of Mnase. In the experiments represented here, the repeat length was calculated from the slope of the linear regression for each lane where there were at least three nucleosomal bands visible (Fig12). It was evident from the gradients that the wildtype is more open to Mnase digestion than the *set1*⁻ strain. There is an observable variation of repeat length between these two strains. At low Mnase concentrations the repeat length is higher and theres a greater variation for this value at higher Mnase concentrations. There is little variation between the sizes for the average repeat length at higher concentrations although the values do reflect cleavage within the nucleosome core.

University of Cape Town

Wildtype



University of Cape Town

Set1⁻

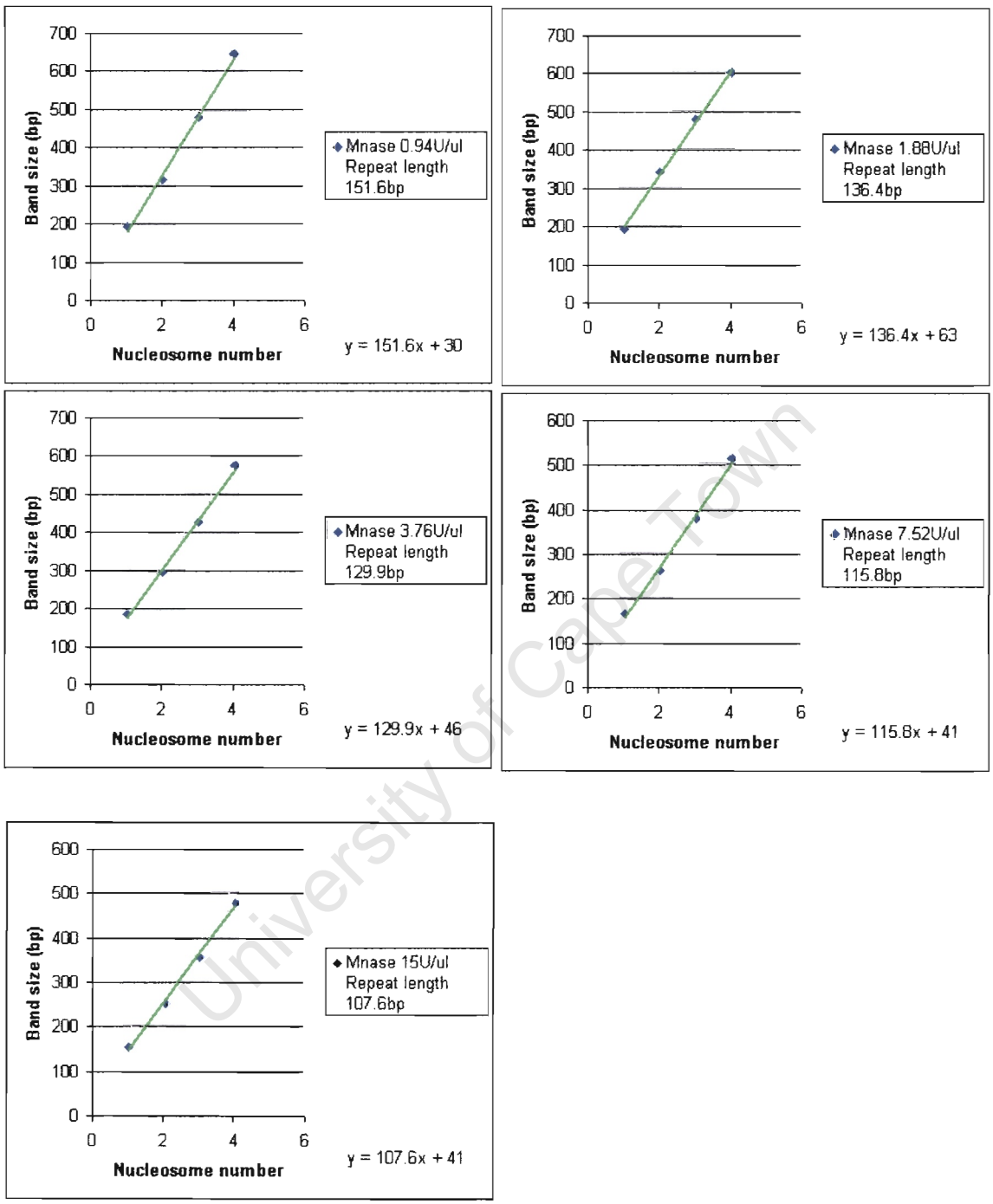


Figure 12. Determinations of nucleosome repeat length for a wildtype and *set1*⁻ strains

The nucleosome repeat lengths were calculated from the slope of the linear regression in a wildtype and a *set1*⁻ strain.

University of Cape Town

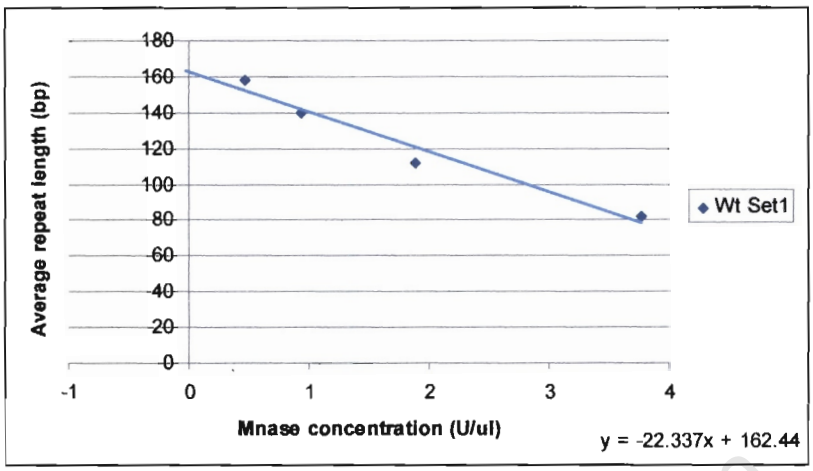
3.2.4 DNA repeat length and Mnase concentration

Since observed nucleosome repeat lengths are sensitive to Mnase concentration; mostly due to sliding at higher Mnase concentrations, the nucleosome repeat length is most accurately determined by extrapolation of the relation between Mnase concentration and apparent repeat length to zero Mnase concentration, as is shown in Figure 13 (van Holde, 1989) .

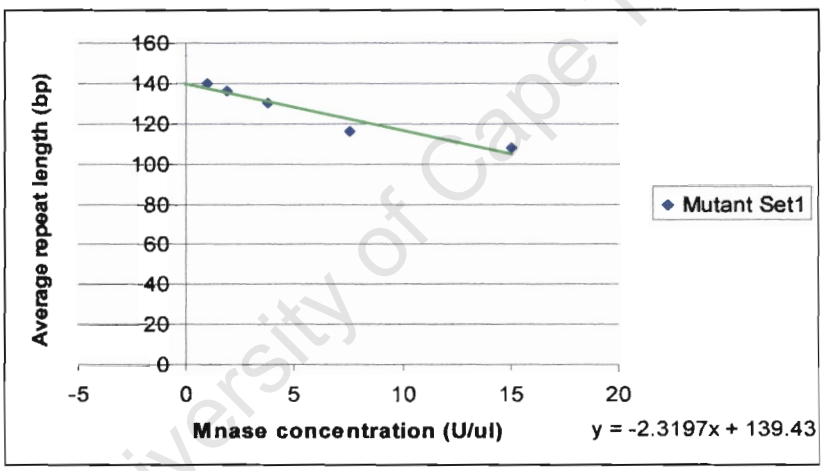
The graphs in Fig 13 depict the change of nucleosome repeats with the concentration of Mnase in the wildtype and *set1*⁻ deleted strains. These repeat lengths changes in both strains with given Mnase concentrations. It can clearly be seen from this graphs that the wildtype has a more rapid trimming of the chromatin. The difference is a graphical confirmation that the chromatin in the wildtype is digested down to smaller sizes more rapidly than the *set1*⁻ strain. The nucleosome repeat length of the *set1*⁻ strain is shorter than the wildtype strain. This is evident by the extrapolation of Mnase concentration equal to zero (Fig 13), which showed that the wildtype has a repeat length of 162.5bp and 139.42bp in the *set1*⁻ strain.

University of Cape Town

Wildtype



Set1⁻



University of Cape Town

Figure 13: DNA repeat lengths as a function of Mnase concentration for a wildtype and *set1*⁻ strains.

The data shown are for a wildtype and a *set1*⁻ strain with increased [0.47-15 u/µl] Mnase concentration. The value at 0U/µl Mnase gives the nucleosome repeat length.

3.3 Summary

Yeast nuclei were successfully isolated from a wildtype and a *set1*⁻ strain. These nuclei were treated with increasing concentrations of Mnase, and the length of the DNA associated with the chromatin fragments analyzed on an agarose gel. From the gel and corresponding graph it was determined that the *set1*⁻ strain was more resistant to Mnase cleavage in comparison to the wildtype strain. From the graph, the repeat lengths were determined for both the wildtype and the *set1*⁻ strain. From these values the wildtype appeared to be more accessible to Mnase in the linker region than the *set1*⁻ strain. The repeat lengths were smaller in the *set1*⁻ strain. This could be clearly seen when the repeat lengths were plotted against their corresponding Mnase concentrations

University of Cape Town

Chapter 4: The positioning of nucleosomes obtained from *in vivo* chromatin digestions and indirect end-labeling at the *TRX2* and *SIR4* promoters

4.1 Introduction

Many advances have been made in understanding the relationship between histone post-translational modifications, chromatin structure and transcription. It was shown by Hui Ng *et al.*, (2003) that the Set1p methyltransferase was targeted to the 5' end of mRNA coding regions via the recognition of RNA polymerase II associated proteins Kin 28, Rtf1 and Paf1.

I have shown by micrococcal nuclease digests of a wildtype and *set1*⁻ deleted strain that in the absence of Set1p methyltransferase, there was an apparent change in chromatin structure as suggested by a change in nucleosome accessibility (Chapter 3). To understand this change in chromatin structure, I chose two candidate gene promoters that displayed a significant increase or decrease in transcriptional activity in the absence of histone H3K4 methylation (Boa *et al.*, 2003). The genes chosen were the *silent information regulator 4* gene *(SIR4)* which had a 2 fold increase, and the *LMA1* or thioredoxin 2 *(TRX2)* gene, which had an 8 fold decrease in transcriptional activity in the absence of H3K4 methylation. In our approach to understand how nucleosomes positions are affected at these promoters, low resolution mapping was performed. We were specifically interested in the position of nucleosomes on the promoters of the *SIR4* and *TRX2* genes, particularly in whether H3K4 methylation dependent differences was

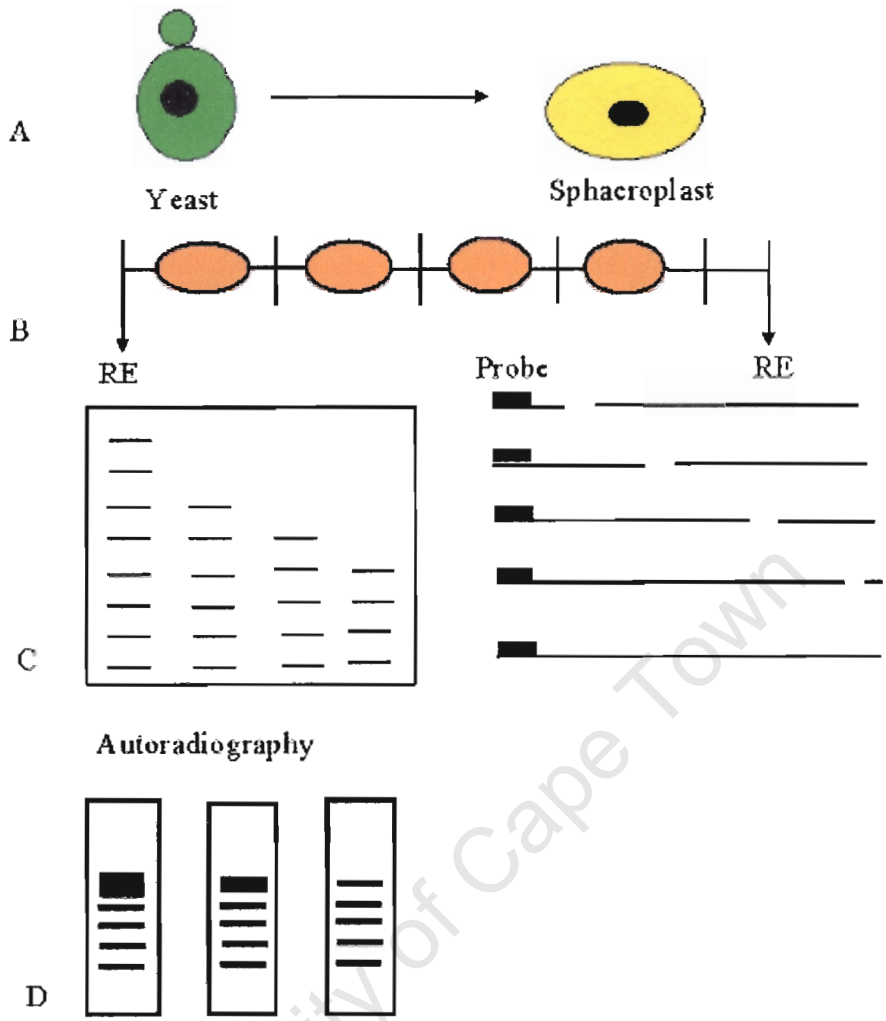
University of Cape Town

present, which would provide a link, possibly indirect, between H3 methylation and gene expression.

4.2 Chromatin organization of the *SIR4* promoter for a wildtype and *set1*⁻ strain *in vivo*.

To determine whether the upregulation of the *SIR4* gene is due to a change in the chromatin structure, sphaeroplasts were isolated from a wildtype and *set1*⁻ strain. Low resolution chromatin analysis was performed by indirect end- labeling in order to analyze two regions within the *SIR4* gene. To map these nucleosomes, sphaeroplasts were isolated from both a wildtype and *set1*⁻ strain (Fig14). These sphaeroplasts were treated with different concentrations of Mnase, and the DNA fragments resolved by agarose gel electrophoresis. For indirect end labeling, the Mnase digested DNA was cut with a restriction endonuclease that cleaved adjacent to the area of interest. The fragments were then run on a large agarose gel, and the gel was blotted to a nitrocellulose membrane. The membrane was probed with a radiolabeled PCR fragment corresponding to an end bordering the cleavage sequence. The indirect end labeling technique was performed adjacent to a DraI restriction site, situated upstream of the *SIR4* promoter.

University of Cape Town



University of Cape Town

Figure 14: The chromatin indirect end labeling method

(A) isolation of yeast sphaeroplasts **(B)** Secondary digest with a restriction enzyme 5'to the probe**(C)** Transfer to a nitrocellulose membrane and probe with radiolabelled PCR product.**(D)** Autoradiography (Hörtz *et al.*,1998).

4.2.1Chromatin structure of the *SIR4* promoter at position -280- to +233

A probe was generated from a *SIR4* gene sequence from position -280 to position +233. This sequence was analyzed for consensus binding elements and transcription factors using the Transfac database. Structural features of the SIR4 promoter include the ATG start codon (+1), the TATA box (-1) and the poly (dA-dT) (-2) sequences (Fig15). The isolated sphaeroplasts from a wildtype and *set1*⁻ strain were digested with increasing amounts of Mnase, digested with DraI and BstEII restriction enzymes, electrophoresed (Fig16), transferred to a nitrocellulose membrane, and probed with the 513bp PCR product which was labeled with [α-^{32}P] dCTP.

From these results (Fig 17) in the *set1*⁻ strain (lanes 7, 8, 9), there appears to be significant difference in the accessibility of the DNA in the linker region compared to the nucleosomally wrapped DNA (lanes 10-13). This is evident when looking at nucleosome core particles NCP -2 and -1. It also appears that the DNA is more extensively cleaved in the *set1*⁻ strain compared to the wildtype strain (lanes 4, 5, 6). There is a slight difference in the locations of cleavage within the linker DNA between the *set1*⁻ and wildtype strains, suggesting that the nucleosomes may be differently positioned on the *SIR4* promoter in the *set1*⁻ and wildtype strains. Since the methylation by Set1p occurred on the H3 N-terminal tails, it is unlikely to have influenced nucleosome positions due to differences in the binding of the centrally located octamer. Although there are clearly defined NCPS at these positions with increasing Mnase, the apparent loss of nucleosome association at position +1 is due to the digestion extent of one sample.

University of Cape Town

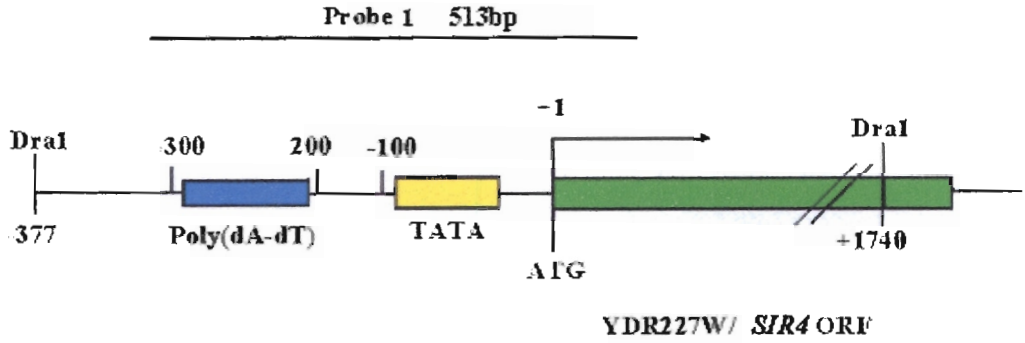
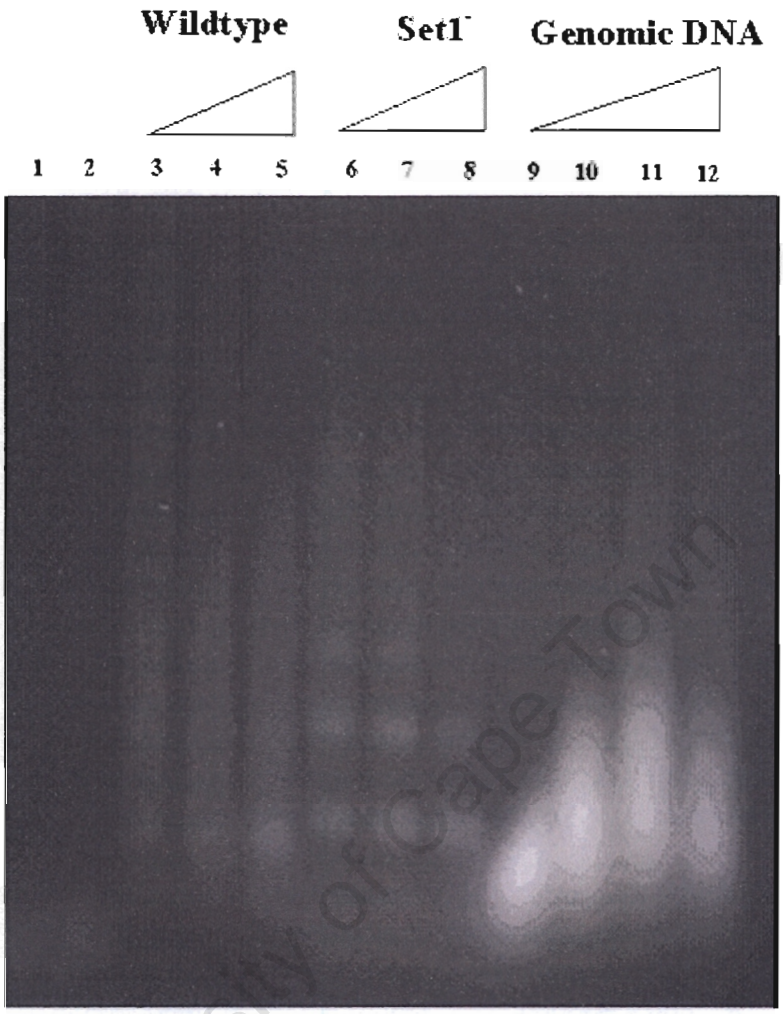


Figure 15: Map of the *SIR4* promoter in the yeast *S. cerevisiae*

The localization of the two probes used in the analysis of the *SIR4* promoter. The functional sites, ATG- start codon, TATA box and a poly (dA-dT) stretch are regions for transcription factor binding sites. The sites which flank the *SIR4* gene are defined by two DraI sites. Probe 1 hybridizes in region -280 to +233.

University of Cape Town



University of Cape Town

Figure 16: Digests of wildtype and *set1*⁻ sphaeroplasts and naked genomic DNA

Genomic DNA was digested with 10U/µl restriction enzymes BstEII (lane 1), DraI (lane2) and DraI/BstEII (lane3). Chromatin from wild type and the *set1*⁻ sphaeroplasts was digested with 0.47U/µl (lane 4), 0.94U/µl (lanes 5 and 7), 1.88U/µl (lanes 6 and 8) and 3.88U/µl (lane 9) Mnase. Genomic DNA was digested with 2.0U/µl (lane10), 0.75U/µl (lane11), 0.20U/µl, (lane 12) and 0.10U/µ1 (lane13) Mnase. The Mnase digests were subjected to secondary digests with 10U/µ1 of DraI / BstEII. This is an ethidium bromide stained agarose gel visualized by UV trans illumination.

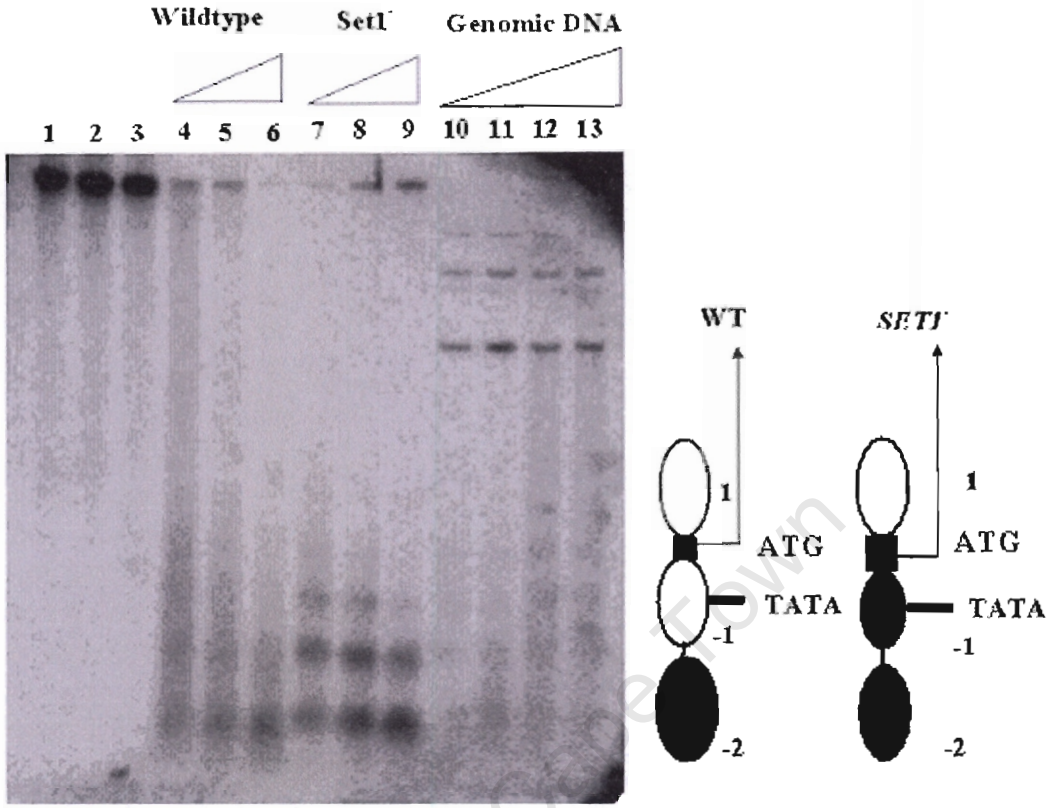


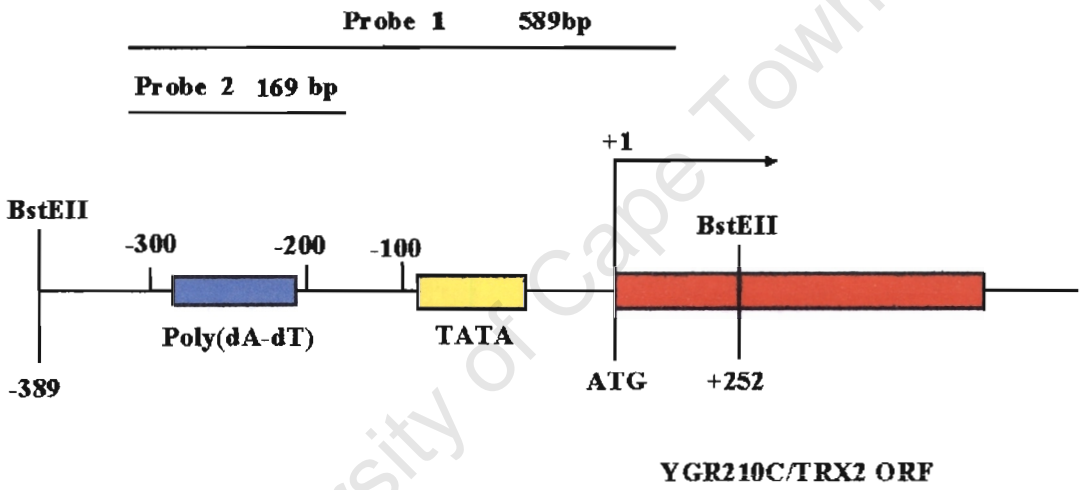
Figure 17: The chromatin organization of the *SIR4* promoter at position -280- to +210 in a wildtype and *set1*⁻ strain during log phase growth.

Mnase cleavage sites were mapped relative to a 513bp probe adjacent to a DraI site. Genomic DNA were digested with 10U/µl restriction enzymes BstEII (lane 1), Dra1 (lane2) and DraI/BstEII (lane3). Chromatin from wild type and *set1*⁻ sphaeroplasts were digested with 0.47U/µl (lane 4), 0.94U/µl (lanes 5 and 7), 1.88U/µl (lanes 6 and 8) and 3.88U/µl (lane 9) Mnase. Genomic DNA was digested with 2.0U/µl (lane10), 0.75U/µ/ (lane11), 0.20U/µl, (lane 12) and 0.10U/µ1 (lane13) MNase. The Mnase digests were subjected to secondary digests with 10U/µ1 of DraI / BstEII. Protected areas that appear to correspond to positioned nucleosomes are shown on the side of the gel. White ovals represent nucleosomes which are not clearly defined and dark ovals represent nucleosomes which are clearly defined. This is a phosphoimager of the hybridized membrane.

University of Cape Town

4.3 Chromatin organization of the *TRX2* promoter for a wildtype and *set1*⁻ strain at low resolution

To determine whether the repression of the *TRX2* protein is due to a change in the chromatin structure we employed the same methods as for the *SIR4* promoter chromatin mapping. This gene had an 8 fold decrease in transcriptional activity as shown by microarray analysis (Boa *et al.*, 2003).



University of Cape Town

Figure 18: Map of the *TRX2* promoter in the yeast S. *cerevisiae*

The localization of the two probes used in the analysis of the *TRX2* promoter. The localization of the two probes used in the analysis of the *TRX2* promoter. The functional sites, ATG- start codon, TATA box and a poly (dA-dT) stretch are regions for transcription factor binding sites. The sites which flank the *TRX2* gene are defined by two BstEII sites. Probe 1 hybridizes in region -339 to +250. Probes 2 hybridize from region -339 to -170.

4.3.1 Chromatin structure of the *TRX2* promoter

The *TRX2* probe was generated from a *TRX2* gene sequence from position -339 to position +250. This sequence was analyzed for consensus binding elements for transcription factors with the Transfac database. Structural features of the *TRX2* promoter are the ATG (+1) start codon, the TATA box (-1) and the poly (dA-dT) (-2) sequences (Fig 18).

The isolated sphaeroplasts from a wild type and *set1*$^-$ strain were digested with increasing amounts of Mnase, digested with DraI and BstEII restriction enzymes, electrophoresed (Fig19), transferred to a nitrocellulose membrane and, probed with the 589bp PCR fragment which was labeled with [α-^{32}P] dCTP.

Fig 20 illustrates the chromatin structure of the *TRX2* promoter and reveals that there are two nucleosome positions, at NCP -2 and -1. The *set1*$^-$ strain has increased accessibility to Mnase in the linker region at NCP -2 compared to the nucleosomally wrapped DNA. The wildtype appears to have similar levels of Mnase accessibility at NCP -2 as in the *set1*$^-$ strain. In Fig 20 the wildtype footprints is slightly upstream compared to the *set1*$^-$ strain, suggesting that at NCP -2 there is a change in chromatin structure or a change in the nucleosome position. NCP- 2 is situated on a possible poly (dA-dT) site and the H3K4 methylation is unlikely to influence the binding of the nucleosome core.

University of Cape Town

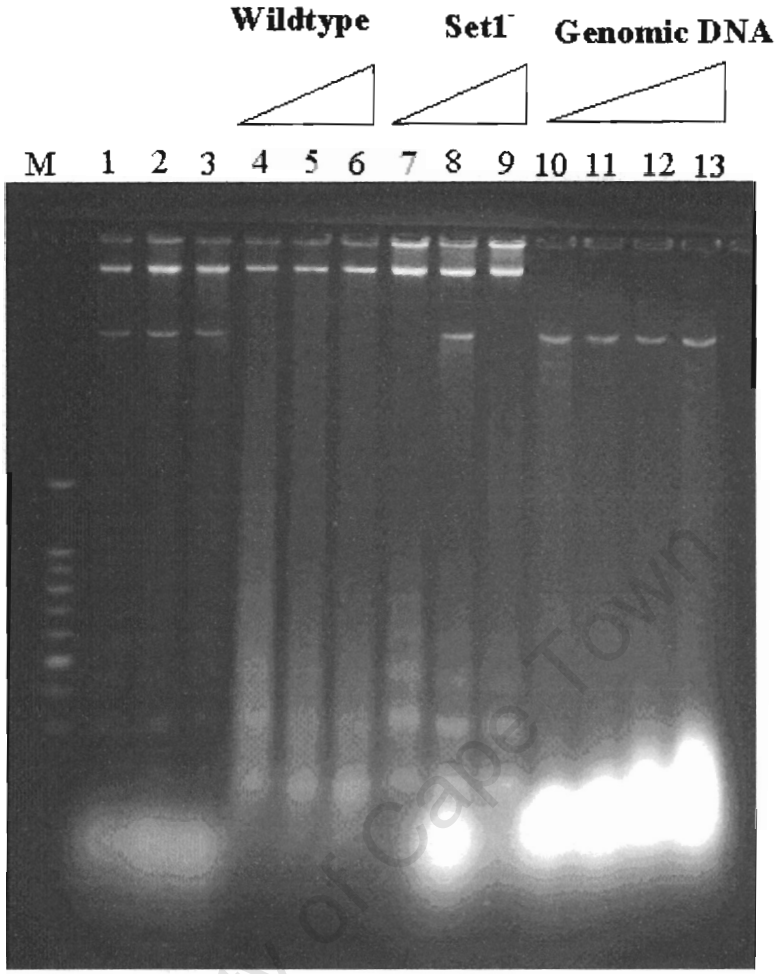


Figure 19: Digests of wildtype and *set1*⁻ sphaeroplasts and naked genomic DNA with increasing concentrations of Mnase.

Marker 100bp followed by genomic DNA which were digested with 10U/µl restriction enzymes BstEII (lane 1), DraI (lane2) and DraI/BstEII (lane3). Chromatin from wild type and the *set1*⁻ sphaeroplasts was digested with 0.47U/µl (lane 4)), 0.94U/µl (lanes 5 and 7), 1.88U/µl (lanes 6 and 8) and 3.88U/µl (lane 9) Mnase. Genomic DNA was digested with 2.0U/µl (lane10), 0.75U/µl (lane11), 0.20U/µl (lane 12) and 0.10U/µ1 (lane13) Mnase. The Mnase digests were subjected to secondary digests with 10U/µ1 of DraI and BstEII. This is an ethidium bromide stained agarose gel visualized by UV trans illumination.

University of Cape Town

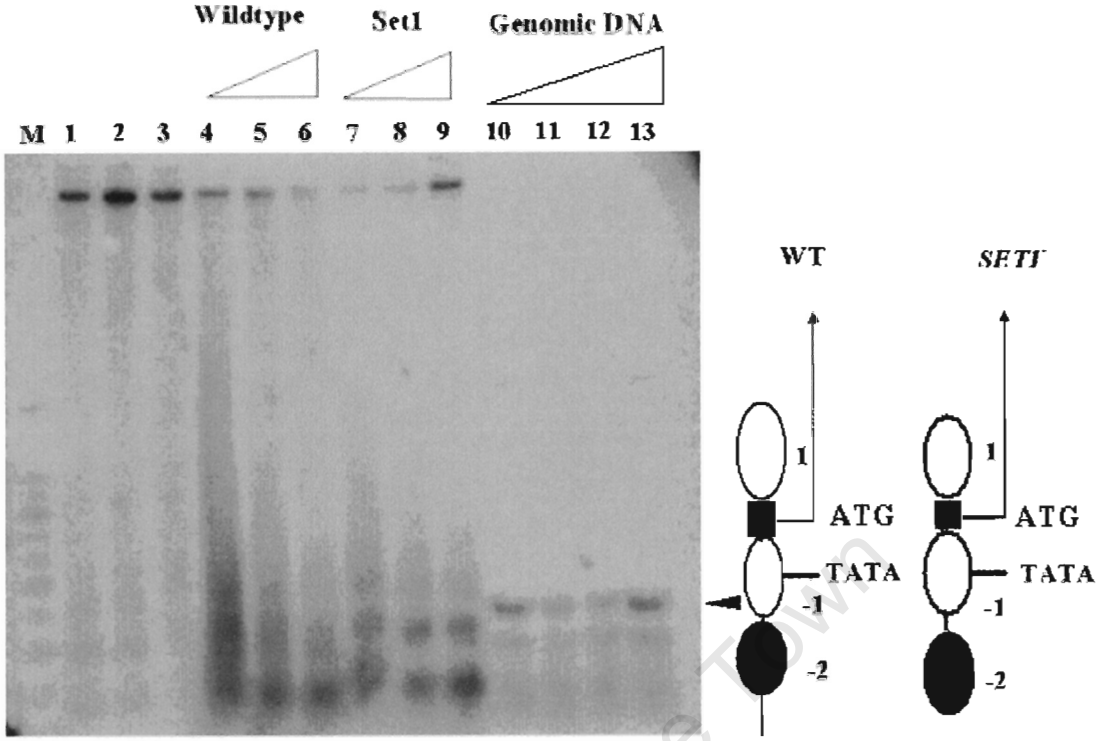


Figure 20: The chromatin organization of the *TRX2* promoter at position -339 to +250 in a wildtype and *set1*⁻ strain during log phase.

Mnase cleavage sites were mapped relative to a 589bp probe adjacent to a BstEII site. Marker 100bp followed by genomic DNA which were digested with 10U/µl restriction enzymes BstEII (lane 1), DraI (lane2) and DraI/BstEII (lane3). Chromatin from wild type and the *set1*⁻ sphaeroplasts was digested with 0.47U/µl (lane 4), 0.94U/µl (lanes 5 and 7), 1.88U/µl (lanes 6 and 8) and 3.88U/µl (lane 9) Mnase. Genomic DNA was digested with 2.0U/µl (lane10), 0.75U/µl (lane11), 0.20U/µl (lane 12) and 0.10U/µ1 (lane13) Mnase. The Mnase digests were subjected to secondary digests with 10U/µ1 of DraI and BstEII. The nucleosome positions are indicated on the side. White ovals represent nucleosomes which are not clearly defined and dark ovals represent nucleosomes which are clearly defined. This is a phosphoimager of the hybridized membrane.

University of Cape Town

4.3.2 Analysis of the *TRX2* promoter at position -339 to -170

The 589bp probe gave an indication of the chromatin organization at the *TRX2* promoter region. These results were confirmed by shortening the 589bp probe (-339 to +250) to 169bp (-339 to -170). The same method as described above was used to analyze this region but the Mnase digest was done for 5 minutes. The genomic DNA digest was done at lower Mnase concentrations as to ensure not to over digest the DNA (Fig 21). The 169bp probe was labeled with [α-^{32}P] dCTP, the blot was hybridized, and exposed to a phosphor screen for two days.

The results obtained were similar to those seen with the 589bp probe (Fig 22). As previously seen in Fig 20, there appeared to be nucleosomes associated at both NCP -1 and -2. With the longer probe the *set1*$^-$ strain was slightly more accessible to Mnase at NCP-2 (Fig 22). The footprints were shifted slightly upstream in the *set1*$^-$ strain in comparison to the wildtype strain, confirming the first result which showed that there was a difference in nucleosome positions.

University of Cape Town

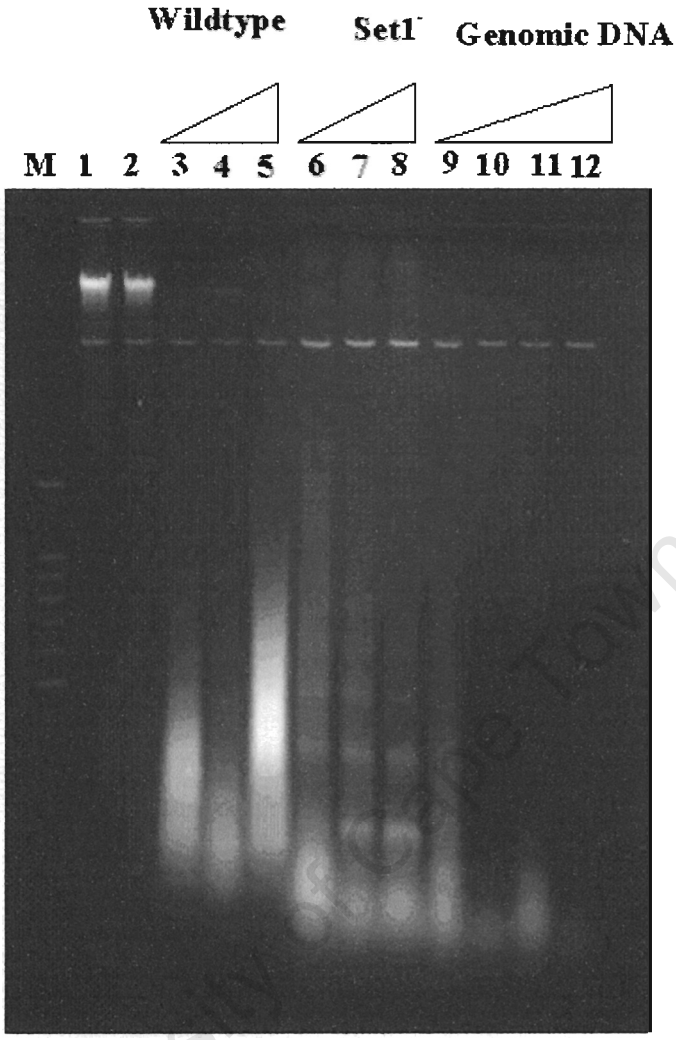
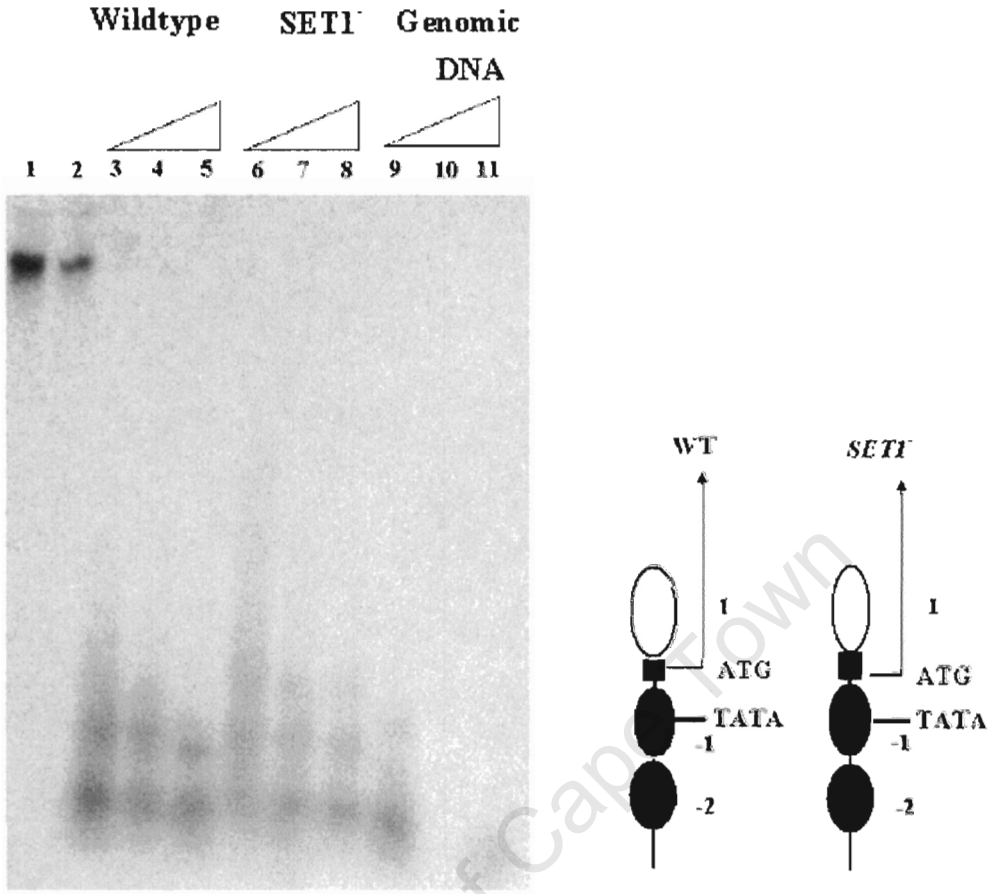


Figure 21: Digests of wildtype and *set1*⁻ sphaeroplasts and naked genomic DNA with increasing concentrations of Mnase.

Genomic DNA restriction enzyme digest were done with 10U/µl BstEII (lane 1) and DraI (lane2) Chromatin from wild type and the *set1*⁻ sphaeroplasts was digested with 0.47U/µl (lane 3), 0.94U/µl (lanes 4 and 6), 1.88U/µl (lanes 5 and 7) and (lane 8) 3.88U/µl Mnase. Genomic DNA was digested with 10U/µl (lane9), 2U/µl (lane10), 1U/µl, (lane 11) and 0.10U/µ1 (lane12) Mnase. The Mnase digests were subjected to secondary digests with 10U/µ1 of DraI and BstEII. This is an ethidium bromide stained agarose gel visualized by UV trans illumination.



University of Cape Town

Figure 22: The chromatin organization of the *TRX2* promoter at position -339 to -170 in a wildtype and *set1⁻* strain during log phase.

Mnase cleavage sites were mapped relative to a 169 bp probe adjacent to a *BstEII* site. Genomic DNA restriction enzyme digest were done with 10U/µl BstEII (lane 1) and Dra1 (lane2) Chromatin from wild type and the *set1⁻* sphaeroplasts was digested with 0.47U/µl (lane 3), 0.94U/µl (lanes 4 and 6), 1.88U/µl (lanes 5 and 7) and 3.88U/µl (lane 8) Mnase. Genomic DNA was digested with 10U/µl (lane9), 2 U/µ/ (lane10) and 1U/µl, (lane 11) Mnase. The Mnase digests were subjected to secondary digests with 10U/µ1 of DraI and BstEII. The nucleosome positions are indicated on the side. White ovals represent nucleosomes which are not clearly defined and dark ovals represent nucleosomes which are clearly defined. This is a phosphoimager of the hybridized membrane.

4.4 Summary

Two promoter regions were analyzed each with two different probe lengths. The larger probe was used for general analysis and the smaller for high resolution mapping. The *SIR4* gene was upregulated and the *TRX2* was downregulated in the *set1*⁻ strain. The low resolution method provided us with a clear insight into the different levels of structural organization during the growth phase of a wildtype and *set1*⁻ strain. The results indicated that there were differences between the chromatin structure and transcription at the *SIR4* promoter.

The high resolution mapping at the *SIR4* promoter showed that the DNA was more extensively digested in the *set1*⁻ strain in comparison to the wildtype. The increased Mnase accessibility at NCP-2 in the *set1*⁻ strain suggested alternate nucleosome positions, longer repeat length or a lower nucleosome density.

The chromatin organization of the *TRX2* promoter at high resolution showed that the footprints were higher in the *set1*⁻ strain than in the wildtype. This suggested that there was a change in nucleosome position.

University of Cape Town

Chapter 5: Discussion

Histone lysine methylation has become synonymous with gene silencing and heterochromatin formation. Biochemical and genetic studies by Nielsen *et al.*, (2001) and Schultz *et al.*, (2002) showed that gene silencing was dependent on the interaction of histone H3K9 with the HP1 protein, a transcriptional repressor.

In contrast, budding yeast *S. cerevisiae* was shown to lack H3K9 methylation. The mechanism by which genes are silenced in *S.cerevisiae* is still unclear, but recently it was shown that H3K4 methylation is present at both active and inactive genes.

In *S.cerevisiae,* the Set1p methyltransferase functions as an 8 protein complex known as COMPASS, which is responsible for 50% of the total H3K4 methylation (Krogan *et al.*, 2002; Boa *et al.*, 2003). Functional studies have shown that this methylation mark is responsible for both ribosomal and telomeric silencing, which is independent of Sir*2*p binding (Bryk *et al.*, 2002).

The hypothesis was that these antagonistic and repressive states could be due to a local change in chromatin structure. I tested this hypothesis through the isolation and digestion of wildtype and *set1*$^{-}$ nuclei and sphaeroplasts from *S.cerevisiae* with micrococcal nuclease. From these results I observed a significant change in nuclease susceptibility between a wildtype and the *set1*$^{-}$ strain. The wildtype was much more sensitive to Mnase cleavage at the linker region in comparison to the *set1*$^{-}$ strain. From these observations it was clear that the *set1*$^{-}$ strain had an appearance more typical of heterochromatin, which, when cleaved with Mnase, produces sharper, more intense nucleosomal bands. It was known that heterochromatin has a reduced linker length (Dillon and Festenstein, 2002).

University of Cape Town

This extra protection at linker regions could be evidence of chromatin binding protein such as the Sir proteins. It was shown that the Sir3p and Sir4p proteins could bind to the tails of histones H3 and H4 (Rusche, *et al.*, 2002). This interaction could cause the spread of Sir3p and Sir4p proteins across the genome. The methylation of H3K4 was shown to be involved in telomeric silencing and, in comparison, the expression of the Sir3 protein influences silencing at telomeres (Bryk *et al.*, 2002; Wolffe, 1998). The Sir3 and Sir4 proteins may not be dependent on methylation of H3K4, but they exert the same characteristics as at silenced loci. This could be one of the reasons why *set1*$^{-}$ Mnase digest has the appearance of heterochromatin, and further evidence showed it had a shorter repeat length. In contrast, histone methylation (mono- and di-methylation) was present at both active and inactive genes. The concentration of a particular level of methylation at these areas could influence Sir3p and Sir4p binding and heterochromatin formation.

The wildtype chromatin could clearly be seen to be sensitive to increased levels of Mnase. The bands were not as distinct as in the *set1*$^{-}$ strain and this corresponds to data by Hui Ng *et al.*, (2003) who showed that the occupancy of Set1p causes an increase in RNA polymerase II transcription. This is evident when looking at the differences in the nucleosomes positioned on both the S*IR4* and *TRX2* promoters. In both cases the footprint of the nucleosomes in the *set1*$^{-}$ strains appeared slightly smaller than in the wildtype strain. This could be due to local changes in chromatin structure. In addition it could also due to various transcription and remodeling factors which are not only dependent on histone H3K4, methylation but also the occupancy of the Set1p protein. One such example is from data obtained from Santos-Rosa *et al.*, (2003), which showed

University of Cape Town

that at certain loci the Isw1p ATPase is dependent on H3K4 methylation at the *Met16* locus. Although in the absence of this modification it did not completely abolish the binding of the Isw1p ATPase (Santos-Rosa *et al.*, 2003). It was thus evident that the *S. cerevisiae* had certain spatial boundaries which are not solely dependent on H3K4 methylation for transcription.

By comparing the repeat lengths it was evident that the *set1*⁻ stain was less susceptible to Mnase cleavage than the wildtype within the linker region. It was known that the lysine rich histones (H1, H5) and non-histone proteins interacted with the linker DNA. During active transcription with the removal of H1, there could be a variation in repeat length between both strains. The average repeat length of the wildtype was shorter than the *set1*⁻ strain for the corresponding Mnase concentrations. In contrast, the *set1*⁻ strain had repeat lengths which were larger than the wildtype, but significantly shorter than the ~160bp calculated for *S. cerevisiae*. The results of the Mnase digest clearly showed that there is a difference between the wildtype and *set1*⁻ chromatin during log phase growth.

University of Cape Town

It was previously reported that in the absence of H3K4 methylation, there was a decrease of approximately 80% of genes (Boa *et al.*, 2003). The loss of methylation of H3K4 by Set1p has both positive and negative effects on particular genes during transcription. It was documented that the presence of nucleosomes was an indication of gene repression, and the dissociation of nucleosomes was associated with transcriptional activation. The results obtained through Mnase digest and indirect end labeling of the *SIR4* promoter, showed that there was a difference in the chromatin structure between the wildtype and *set1*⁻ strain. The microarray data showed that there was a twofold increase in the

transcription of the Sir4 protein. The expected result would have been a loss of nucleosomes which would be an indication of transcriptional activation. In contrast, nucleosomes could clearly be seen at positioned at NCP -1 and -2. There was one nuclease-hypersensitive site at the promoter region corresponding to the poly (dA-dT) region. This hypersensitive site was clearly more defined in the *set1*⁻ strain than in the wildtype. This could be a possible *cis*–acting sequence which would regulate the expression of the *SIR4* gene by *trans*- acting factors. This would be one of the reasons why there was a two fold increase in *SIR4* expression in the *set1*⁻ strain. It was shown by Hui Ng *et al.*, (2003) that the Set1 is recruited to the 5'end by RNA polymerase II elongation machinery which would lead to nucleosome dissociation. The *set1*⁻ strain had no nucleosome associated at NCP +1. The methylation of H3K4 at NCP +1 could be required for chromatin remodeling, and the regulation of the *SIR4* gene, but not necessary for RNA polymerase II elongation. The loss of Set1 interaction with the RNA polymerase II elongation machinery could influence gene transcription in a stoichiometric manner, which would lead to a reduction in the level RNA polymerase II or one of its subunits (Boa *et al.*,2003).

The *TRX2* gene had an 8 fold decrease in transcription, which was experimentally determined from microarray analysis (Boa *et al.*, 2003). One Mnase hypersensitive site was detected at NCP-2 in both the wildtype and the *set1*⁻ strain. Both the wildtype and *set1*⁻ strain footprints were similar, as both had nucleosomes positioned at NCP -2. The only difference was the footprints in the *set1*⁻ strain were slightly higher at NCP -2 in comparison to the wildtype. The reason for this could be that the *trans* -acting factors

University of Cape Town

could not gain access to the underlying regulatory sequence which is dependent on the H3K4 methylation for transcription.

University of Cape Town

Chapter 6: Conclusion

The chromatin structure of a wildtype and *set1*⁻ strain was analyzed. The chromatin of the wildtype was more prone to Mnase cleavage than that of the *set1*⁻ strain, which exhibited the appearance of heterochromatin. The high resolution mapping of the *SIR4* and *TRX2* promoters showed that there was a significant difference in the nucleosomal organization in the *set1*⁻ strain in comparison to the wildtype strain. These differences were attributed to a change in nucleosomal position or differences in the repeat length. These findings suggested that the *S. cerevisiae* Set1p caused a change in the chromatin structure by methylating H3K4. The mechanism by which this modification influences chromatin structure is still unclear.

There have been various papers describing the mechanism by which H3K4 methylation by Set1p could influence the chromatin structure. It was shown that the different methylation states of H3K4 was important for a genes transcriptional state and this contributes to the association of Isw1p ATPase with chromatin on certain yeast genes (Santos-Rosa *et al.*,2003).

The Set1p localization at the 5' end of ORFs was shown to be dependent on the TFIIH-associated kinase that phosphorylates the Pol II C- terminal domain and is responsible for the transition between initiation and elongation (Hui Ng *et al.*, 2003).

These findings suggest that the factors which influence chromatin structure are dependent H3K4 methylation. Various other histone post-translational modifications as well as other

University of Cape Town

chromatin binding proteins which are dependent on these modifications could also influence the chromatin structure.

From these experiments there was a clear link between H3K4 methylation by Set1p and the differences between chromatin organization within *S. cerevisiae*. The gene loci that were analyzed had a significant change in chromatin structure which may not be directly influenced by H3K4 methylation. The levels of methylation at H3K4 in certain parts of the genome may act as boundaries which prevented the spreading of proteins involved in gene silencing. The position of these gene loci as well as their methylation states could be factors attributing to their expression in the absence of Set1 methylation of H3K4.This methylation mark may influence the functions of various transcription and remodeling factors.

Future experiments making use of chromatin immunoprecipitation (Chip) techniques, by directing antibodies against di- and tri- methylated H3K4 at gene loci known to be influenced by these methylation marks. This technique can be used to understand the long term effects of this methylation mark as well as its propagation throughout the cell cycle. Experiments which could be useful is the site directed mutagenesis of amino acid residues within the active site involved in the methyl transfer to H3K4. Further microarray and structural analysis should be done to determine to what degree the individual methylation states have on the *Saccharomyces cerevisiae* genome.

University of Cape Town

References

1. Adams, A., Gottschling, D.E., Kaiser, C.A. and Stearns, T. (1998) Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual. Cold Spring Harbor Laboratory Press, Plainview, New York

2. Ausubel, F., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1995) Short Protocols in Molecular Biology. John Wiley & Sons, Inc, Boston.

3. Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C. and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, **410**, 120-124

4. Beisel, C., Imhof, A., Greene, J., Krimmer, E. and Sauer, F. (2002). Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1 *Nature*, **419**, 857-862

5. Beuchle, D., Struhl, G. and Müller, J. (2001). Polycomb group proteins and heritable silencing of *Drosophila* Hox genes. *Development*, **128**, 993-1004

6. Berger, S.L. (2001). The histone modification circus. *Science*, **292**, 64-65

University of Cape Town

7. Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T. and Schreiber, S.L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc.Natl.Acad.Sci. USA*, **99**, 8695-8700

8. Birve, A., Segupta, A.K., Beuchle, D., Larsson, J., Kennison, J.A., Rasmuson-Lestander and Müller, J. (2001). Su(z)12, a novel *Drosophila* Polycomb group gene that is conserved in vertebrates and plants. *Development*, **128**, 3371-1179

9. Boa, S., Coert, C. and Patterton, H.G. (2003). *Saccharomyces cerevisiae* Set1p is a methyltransferase specific for lysine 4 of histone H3 and is required for efficient gene expression. *Yeast*, **20**, 827-835

10. Briggs, S.D., Bryk, M., Strahl. B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* **15**, 3286-3295

11. Bryk, M., Briggs, S.D., Strahl, B.D., Curcio, M.J., Allis, C.D. and Winston, F. (2002). Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in *S. cerevisiae* by a Sir2- independent mechanism. *Curr. Biol.* **12**, 165-170.

University of Cape Town

12. Byrd, K.N. and Shearn, A. (2003). Ash, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci. USA*, **100**, 11535-11540

13. Cao R, Wang L, Wang, H, Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*, **298**, 1039-43.

14. Carvin, C.D. and Kladde, M.P. (2004). Effectors of lysine 4 methylation of histone H3 in *Saccharomyces cerevisiae* are negative regulators of *PHO5* and *GAL1-10*. *J. Biol. Chem*. **279**, 33057-33062.

15. Chantalat, L., Nicholson, J.M., Lambert, S.J., Reid, A.J., Donovan, M.J., Reynolds, C.D., Wood, C.M. and Baldwin, J.P. (2003). Structure of the histone-core octamer in KCl/phosphate crystals at 2.15 Å resolution. *Acta. Cryst*. **59**, 1395-1407

16. Czermin, B., Melfi, R., Mcabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002). *Drosophila* enhancers of zeste/ESC complexes have a histone H3 methyltransferase activity that marks Chromosomal Polycomb sites. *Cell*, **111**,185-96

University of Cape Town

17. Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates J.R. 3rd and Grant PA. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J. Biol. Chem.* **279**, 1867-71.

18. Dillon, N. and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Gen.* **18**, 252-258

19. Dutnall, R.N. and Ramakrishnan, V. (1997). Twists and turns of the nucleosome: tails without ends. *Structure*, **5**, 1255-1259

20. Eskeland, R., Clerking, B., Boeke, J., Bonaldi, T., Regula, J.T. and Imhof, A. (2004). The N-Terminus of Drosophila SU (VAR) 3-9 mediates dimerization and regulates its methyltransferase activity. *Biochemistry*, **43**, 3740-3749

21. Gary, J.D. and Clarke, S. (1998). RNA protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.* **61**, 65-131

22. Gregory, P.D., Barbaric, S. and Hörtz, W. (1998). Analyzing chromatin structure and transcription factor binding in yeast. *Methods: A companion to Methods in Enzymology*. **15**, 295-302

University of Cape Town

23. Gross, D.S., Szent-Gyorgyi and William, T.G. (1986). Yeast as a model system to dissect the relationship between chromatin structure and gene expression. In *Yeast Cell Biology*. (Hicks, J., ed.) pp 345-366, UCLA Symposia on Molecular and Cellular Biology 33

24. Hui Ng, H., Ciccone, D.N., Morshead, K.B., Oettinger, M.A. and Struhl, K. (2003). Lysine 79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian Cells: A potential mechanism for position effect variegation. *Proc.Natl.Acad.Sci. USA*., **100**, 1820-1825

25. Hui Ng, H., Robert, F., Young, R.A. and Struhl, K. (2003). Targeted recruitment of set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell*. **11**, 709-719

26. Jackson, J.P., Lindroth, A.M., Cao, X. and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature*. **416**,556-60

27. Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science*. **293**, 1074-1079

University of Cape Town

28. Kanoh, J., Francesconi, S., Collura, A., Schramke, V., Ishikawa, F., Baldacci, G. and Vincent, G. (2003). The fission yeast spSet1p is a histone methyltransferase that functions in telomere maintenance and DNA repair in an ATM Kinase Rad3-dependent pathway. *J.Mol.Biol.* **326**, 1081-1094

29. Kent, N.A., Bird, L.E. and Mellor, J. (1993). Chromatin analysis in yeast using NP-40 permeabilised sphaeroplasts. *Nucleic. Acid Res.* **21,** 4653-4654

30. Kouzarides, T. (2002). Histone methylation in transcriptional control. *Curr. Opin. Gen. Dev.* **12**, 198-209

31. Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., Shilatifard, Buratowski, S. and Greenblatt, J. (2002). Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **23**, 4207-4218

32. Kwon, T., Chang, J.H., Kwak, E., Lee, C.W., Joachimiak, A., Kim, Y.C., Lee, J.W. and Cho, Y.(2003). Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet. *EMBO.* **22**, 292-303

University of Cape Town

33. Lachner, M., O'Caroll, Rea, S., Mechtler, K. and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*. **410**, 116-120

34. Lachner, M. and Jenuwein, T. (2002). The many faces of histone lysine methylation. *Curr. Opin. Cell. Biol*. **14**, 286-298

35. Lacoste N., Utley R.T., Hunter J.M., Poirier G.G. and Cote J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. *J Biol Chem*. **277**, 30421-30424.

36. Lajeunesse, D. and Shearn, A. (1996). *E* (z): a Polycomb group gene or a trithorax group gene? *Development*, **122**, 2189-2197

37. Liang, G., Lin., J.C.Y., Wei, W., Yoo, C., Cheng, J.C., Weisenberger, D.J., Egger, G., Takai, D., Gonzales, F.A. and Jones, P.A. (2004). Distinct localization of H3 acetylation and H3K4 methylation to the transcriptional start sites in the human genome. *Proc. Natl. Acad. Sci,* **101**, 7357-7362

38. Litt, M.D., Simpson, M., Gasner, M., Allis, C.D. and Felsenfeld, G. (2001). Correlation between histone lysine developmental changes at the chicken β-Globin locus. *Science*, **293**, 2453-2455

University of Cape Town

39. Lee, J. & Bedford, M.T. (2002). PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep.* **3**, 268-73.

40. Lewin, B. (1996). Genes V. Oxford University Press, Oxford

41. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* **389**, 251-260.

42. Manzur, K.L., Farooq, A., Zeng, L., Plotnikova, O., Koch, A.W., Sachchidanand and Zhou, M.M. (2003). A dimeric viral SET domain methyltransferase specific to Lys 27 of histone H3. *Nat. Struct. Biol.* **10**, 187-196.

43. Mathews, C.K. and Van Holde K.E. (1996). Biochemistry 2nd Ed. Bejamin/ Cummings Publishing Company. Inc. Menlo Park, CA.

44. Min J., Zhang X., Cheng X., Grewal, S.I. and Xu R.M. (2002). Structure of the SET domain histone lysine methyltransferase Clr4. *Nat Struct Biol.* **9**, 828-832.

University of Cape Town

45. Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnson, M., Greenblatt, J.F. and Shilatifard, A. (2001). COMPASS: A complex of proteins associated with a trithorax –related SET domain protein. *Proc.Natl.Acad.Sci.USA*, **98**, 12902-12907

46. Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E. & Simon, J.A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell*, **111**,197-208

47. Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. & Grewal, S.I.S. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*, **292**, 110-113

48. Nielsen, S.J., Schneider, R., Bauer, U., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E. and Kouzarides, T. (2001). Rb targets histone H3 methylation and HP1 to promoters. *Nature*, **412**, 561-565

University of Cape Town

49. Nioshioka, K., Chuikov, S., Sarma, K., Erjument-Bromage, H., Allis, C.D., Tempst, P. &Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin 9 formation. *Genes & Dev*. **16**, 479-48

50. Nislow, C., Ray, E. and Pillus, L. (1997). Set1, A yeast member of the Trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol.Biol.Cell*. **8**, 2421-2436

51. Noma, K. and Grewal, S.I.S. (2002). Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. *Proc. Natl. Acad. Sci*. **99**, 16438-16445

52. Nowak, S.J. and Corces, V.G.(2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet*. **20**,

53. Ono, K., Kusano, O., Shimotakahara, S., Mitsuhiro, S., Yamahazi, T. and Shindo, H. (2003). The linker histone homolog Hho 1p from *Saccharomyces cerevisiae* represents a winged helix-turn-helix fold as determined by NMR spectroscopy. *Nucleic Acid Res*. **31**, 7199-7207

University of Cape Town

54. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B., Sun, Z., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, **406**, 593-599

55. Rice, J.C. and Allis, C.D. (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr. Opin. Cell. Biol.* **13,** 263-273

56. Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. and Klug, A. (1984). Structure of the nucleosome core particle at 7Å resolution. *Nature*, **311**, 532-537

57. Richmond, T.J., Searles, M.A. and Simpson, R.T. (1988). Crystals of a nucleosome core particle containing defined sequence DNA. *J. Mol. Biol.* **199**,161-170

58. Richmond, T.J. and Davey, C.A. (2003). The structure of DNA in the nucleosome core. *Nature*, **423**, 145-150

59. Roguev, A., Schaft, D., Shevchenko, A., Pim Pijnappel, W.W.M., Wilm, M., Aasland, R. and Stewart, F. (2001).The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO. J.* **20**, 7137-7148

University of Cape Town

60. Rusche, L.N., Kirchmaier, A.N. and Rine, J. (2002). Ordered nucleation and spreading of silenced chromatin in *Saccharomyces cerevisiae*. *Mol.Biol.Cell*. **13**, 2207-2222

61. Santos-Rosa, H., Schneider, R., Bannister, A. J., Sheriff, J., Berstein, B.E., Tolga Emre, N.C., Schreiber, S.L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature*. **419**, 407- 411

62. Santos-Rosa, H., Schneider, R., Bernstein, B.E., Karabetsou, N., Morillon, A., Weise, C., Schreiber, S.L., Mellor, J. and Kouzarides, T. (2003). Methylation of histone H3 K4 mediates association of the Iswip ATPase with chromatin. *Mol. Cell*. **12**, 1325-1332

63. Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffman, J., Rea, S., Jenuwein, T., Dorn, R. and Reuter, G. (2002). Central role of Drosophila Su (VAR) 3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J.* **21**, 1121-1131

64. Schultz, D.C., Ayanathan, K., Negorev, D., Maul, G.G. and Rauscher III, F.J. (2002). SETDB1: a novel KAP-1 associated histone H3, lysine 9-specific methyltransferase that contributes to HP1- mediated silencing of euchromatic genes by KRAB zinc – finger proteins. *Genes & Dev*. **16**, 919-932

University of Cape Town

65. Sims III, R.J., Nishioka, K. and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet*. **19**, 629-639

66. Simon, J.A. and Tamkun, J.W. (2002). Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. *Curr. Opin. Gen. Dev*. **12**, 210-218

67. Simpson, R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry*, **17**, 5524-5531.

68. Sun, Z. and Allis, D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature*, **418**, 104-108.

69. Strahl, B.D. and Allis, C.D. (2000). The language of the covalent histone modifications. *Nature*, **403**, 41-45

70. Tachibana, M., Sugimoto, K., Fukushima, T. and Shinkai, Y. (2001). Set domain - containing protein, G9a is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem*. **276**, 25309-25317

University of Cape Town

71. Tamaru, H. and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa. Nature,* **414**, 277-283

72. Van Holde, K.E., Allen, J.R., Corden, J., Lohr, D., Tatchell, K. and Weishet, W.O. (1979b). Nuclease digestion and the structure of chromatin. In *Chromatin structure and function Part B: Levels of organization and cell function.* Nicolini, C.A. (ed.) Plenum Press, New York

73. Van Holde, K.E. (1989). Chromatin. Springer Verlag, New York

74. Vaute, O., Nicolas, E., Vandel, L. and Trouche, D. (2002). Functional and physical interaction between the histone methyltransferase Suv38H1 and histone deacetylase. *Nucleic. Acid Res.* **30**, 475-481

75. Warren L. DeLano "The PyMOL Molecular Graphics System."
DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org

76. Wilson, J.R., Jing, C., Walker, P.A., Martin, S.R., Howell, S.A., Blackburn, G.M., Gamblin, S.J. and Xiao, B. (2002). Crystal structure and functional analysis of the histone methyltransferase SET7/9. *Cell.* **111**,105-116

University of Cape Town

77. Wolffe, A.P. (1998). Chromatin: structure and function. Academic Press, San Diego

78. Xiao, B., Jing, C., Wilson, J.R., Walker, P.A., Vaslcht, Kelly, G., Howell, S., Taylor, I.A., Blackburn, G.M. and Gamblin, S.J. (2003). Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature*. **421**, 652-656

79. Xiao, B., Wilson, J.R. and Gambling, S. (2003). Set domains and histone methylation. *Curr. Opin. Struct. Biol.* **13**, 699-705.

80. Zhang, Y. and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the histone tails. *Genes & Dev*. **15**, 2343-2360

81. Zhang, X., Tamaru, H., Khan, S.I., Horton, J.R., Keefe, L.J., Selker, E.U. and Cheng, X. (2002). Structure of the *Neurospora* SET domain protein DIM-5, a histone H3 lysine methyltransferase. *Cell*. **111**, 117-127

82. Zhang L., Eugeni E.E., Parthun M.R. and Freitas M.A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma*. **112**, 77 -86.

University of Cape Town

83. Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U. and Cheng, X. (2003). Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell.* **12**, 177-185

84. Zhang, W., Hayashizaki, Y. and Kone, B.C. (2004). Structure and regulation of the mDot1 gene, a mouse histone H3 methyltransferase. *Biochem. J.* **377**, 641-651

University of Cape Town